

**ASSESSING AFFECTIVE ELEMENTS IN
NEW ZEALAND SECONDARY SCHOOL GENERAL MUSIC
EDUCATION: THE DEVELOPMENT OF A MUSIC ATTITUDE
ASSESSMENT INSTRUMENT BASED ON A TAXONOMY OF
AFFECTIVE EDUCATIONAL OBJECTIVES**

**A Thesis
submitted in partial fulfilment
of the requirements for the
Degree of
Doctor of Philosophy
in the
University of Canterbury**

by

Nicki A. Page

**University of Canterbury
1993**

MT
3
.N5
.P133
1993

CONTENTS

ABSTRACT..... i

ACKNOWLEDGEMENTSiii

PART ONE.....1

CHAPTER I - INTRODUCTION1

 1. GENERAL TOPIC AREA.....1

 2. IMPORTANCE OF THE TOPIC.....23

CHAPTER II - PURPOSE OF THE STUDY.....32

 1 ASSESSMENT OF ELEMENTS IN THE AFFECTIVE DOMAIN -
 AN INTRODUCTION32

 2. NEW DIRECTIONS FOR THE ASSESSMENT OF ELEMENTS IN
 THE AFFECTIVE DOMAIN.....41

CHAPTER III - THE AFFECTIVE COMPONENT OF NEW ZEALAND
SECONDARY SCHOOL GENERAL MUSIC EDUCATION.....52

 1. THE SECONDARY SCHOOL IN NEW ZEALAND.....52

 2. THE DEVELOPMENT OF CONTEMPORARY NEW ZEALAND
 MUSIC EDUCATION59

 3. THE NEW ZEALAND MUSIC EDUCATION REVIEW.....62

 4. THE SYLLABUS FOR SCHOOLS. MUSIC EDUCATION - EARLY
 CHILDHOOD TO FORM SEVEN.....65

CHAPTER IV - REVIEW OF THE RESEARCH.....71

 1. REVIEWS OF RESEARCH INTO ASSESSMENT IN MUSIC.....71

 2. AFFECTIVE DOMAIN AND ASSESSMENT IN MUSIC.....72

 3. STUDIES RELATED TO THE AFFECTIVE TAXONOMY OF
 KRATHWOHL, BLOOM AND MASIA.....88

 4. SUMMARY102

PART TWO104

CHAPTER V - A NEW MUSIC ATTITUDE ASSESSMENT INSTRUMENT104

1. PILOT STUDY	104
2. DEVELOPMENT OF THE NMAAI	112
3. TESTING GROUPS.....	132
4. RESPONDENT INFORMATION SHEET.....	141
CHAPTER VI - RESULTS OF THE NMAAI TRIALS.....	143
1. BACKGROUND	143
2. RESULTS	147
3. SCALOGRAM ANALYSIS.....	165
CHAPTER VII - CONCLUSION: INTERPRETATIONS AND IMPLICATIONS	173
1. SUMMARY OF THE INTRODUCTORY SECTION (Chapters I to IV).....	173
2. DISCUSSION AND INTERPRETATION OF THE NMAAI TRIALS (Chapters V and VI).....	176
3. IMPLICATIONS FOR SECONDARY SCHOOL GENERAL MUSIC EDUCATION.....	187
4. IMPLICATIONS FOR FURTHER RESEARCH	192
Appendix A: The Five Categories of the Affective Taxonomy.....	196
Appendix B: Hevner Instrument.....	197
Appendix C: Noble Instrument	201
Appendix D: Item Level Placement for Noble Instrument	211
Appendix E: Lewy Instrument.....	212
Appendix F: Item Level Placement for Lewy Instrument.....	219
Appendix G: Adapted Lewy Instrument	220
Appendix H: Item Level Placement for Adapted Lewy Instrument	224
Appendix I: The NMAAI	225
Appendix J: Item Level Placement for the NMAAI.....	231
Appendix K: Respondent Information Sheet for the NMAAI.....	232
Appendix L: Summaries for Individual NMAAI Items: Test-Retest Group.....	233
BIBLIOGRAPHY	234

ABSTRACT

The purpose of this study was:

- (1) To examine the issue of assessment in the affective domain, with particular reference to New Zealand secondary school general music education; and
- (2) To make a practical contribution to the assessment of elements in the affective domain, by addressing problems and concerns raised by the examination of affective domain issues.

The affective domain was described as incorporating “positive and negative feelings, as well as emotionally toned attitudes, values, interests and appreciations ...”. It was suggested that too often there is a wide discrepancy between stated affective objectives for a particular course of learning, and any subsequent evaluation of that course. Specific problems and issues were identified, and these were to form a linking thread throughout the thesis.

The taxonomy of affective objectives, developed by Krathwohl, Bloom and Masia, was introduced and discussed, with particular attention being drawn to the position of ‘attitude’ on the taxonomy continuum. The proposal was to base a music attitude assessment instrument on the model: this forming the “practical contribution” component of the study.

A pilot study was undertaken using a previously developed, affective taxonomy-based assessment instrument; the purpose being to gain some indication as to the viability of the proposed project.

Using carefully formulated item statements, a New Music Attitude Assessment Instrument (the NMAAI) was constructed and administered to students in New Zealand secondary schools. A particular characteristic of the NMAAI is that, unlike previous Instruments designed according to similar principles, its underlying definition of 'music' is nonspecific.

The resultant data were analysed, and validity and reliability studies were undertaken.

The "specific problems and issues" introduced in the opening chapters of the thesis were revisited in the light of the NMAAI project, and the findings were detailed.

It was concluded that the NMAAI, as an instrument for use in secondary school general music education, contains demonstrably beneficial properties.

It has the potential to aid the assessment of objectives from the New Zealand music education syllabus for schools.

Implications for further research were identified, both with regard to the NMAAI itself, and for affective domain assessment in general.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, David Sell (School of Music, University of Canterbury), to my parents, Barbara and Ian Short, and to my husband, Hugh Page, for their guidance and unfailing support during the preparation of this thesis.

I would also like to thank Dr Cedric Hall (Victoria University of Wellington) for his expert and timely advice, offered so generously.

PART ONE

CHAPTER I - INTRODUCTION

1. GENERAL TOPIC AREA

The issue of assessment in education - as evidenced by the literature - is certainly no stranger to criticism and controversy. Assessment procedures are continually being reviewed and updated in an effort to maintain an appropriate relationship between them and the policies and practices of modern curricula.

Assessment developments in the field of Music Education are no exception to this effort. Music Educators, in seeking to relate their subject to the wider educational setting, strive to respond to universal advancements and make the seemingly required improvements and modifications.

While this in itself is an admirable, and indeed necessary task, the emphases of many general educational assessment procedures - for example, those linked to cognitive achievement and skill development - are not suited to some aspects of music education. Of particular concern in this respect is the issue of assessing affective domain elements engendered by the subject. A description offered by Thomas Ringness in his book "The Affective Domain in Education" is appropriate at this point:

*The affective domain contains the feelings or emotional components of our lives. Thus positive and negative feelings, as well as emotionally toned attitudes, values, interests and appreciations, morals and character, and even personal and social adjustment fall within that domain.*¹

Assessing these affective domain elements - they being 'subject-specific' - requires special attention from Music Educators, yet they are all too frequently glossed over in deference to

¹ T. Ringness, The Affective Domain in Education, (1975), p. xi.

the above-mentioned 'general assessment emphases'. The reason for the concern regarding this issue is that not enough assistance is offered to Music Educators in the form of constructive, relevant research or theory. Indeed, in many past general music education texts [as typically available in New Zealand] the entire topic of evaluation receives at most one small chapter, generally near the end, as though included as a token gesture. The following short survey illustrates this:

1. Music Education in Action (R.V.D.Morgan and H.N.Morgan):² There is no specific chapter on assessment or evaluation.
 "Tests and Measures" (cognitive only) (pp.162-166) are mentioned in a section entitled "Scientific Devices".
2. Music Education in the High School (Joseph A. Leeder and William S. Haynie):³ Reference is made to assessment under a subheading, "Testing and Reporting" (pp.226-231). Despite the discussion being only five pages long, the authors at least mention the value of assessing "attitude and effort as well as musical progress."⁴ They suggest that "...it is most desirable to devise a marking system which does not spell out complete "failure" in music if the student has a good attitude and shows real effort."⁵
3. Music Education For Teenagers (W.R.Sur and C.F.Schuller):⁶ There is no reference to assessment or evaluation.

² R. Morgan and H. Morgan, Music Education in Action, (1960)

³ J. Leeder and W. Haynie, Music Education in the High School, (1965).

⁴ Ibid, p. 227.

⁵ Ibid.

⁶ W. Sur and C. Schuller, Music Education for Teenagers, 2nd ed., (1966).

4. Music And Young Children (Frances Webber Aronoff):⁷ The seventh chapter, "Evaluation of Music Learning" (pp.149-161), refers to "Affective Objectives" on two pages (149-150); the remainder of the chapter concentrates on cognitive aspects only.
5. Music Curriculum And Instruction (F. Churchley):⁸ The topic of evaluation spans pages 181-185. Two paragraphs only are devoted to "Informal Evaluation", the essence of which may be summed up by the sentence "assessing [receptive attitudes] can be done to a certain extent by informally observing students, both in and out of class."⁹
6. Junior High School General Music (F.M.Andrews):¹⁰ Only four pages (pp. 31-34) are devoted to the topic of evaluation. The assessment of affective objectives is not mentioned.

⁷ F. Aronoff, Music and Young Children, (1969).

⁸ F. Churchley, Music Curriculum and Instruction, (1969).

⁹ Ibid, p. 184.

¹⁰ F. Andrews, Junior High School Music, (1971).

7. Foundations And Principles Of Music Education (Charles Leonhard and Robert House):¹¹ The final chapter, "Evaluation" spans thirty-three pages (pp 389-422), of which fourteen are devoted to the "Evaluation of Students". The only references to affective assessment appear on pages 406-407 under the heading "Evaluating Attitudes", and pages 411-412 under "Evaluating Musical Initiative". With the former, the authors write: "The most common instrument for measuring attitudes is the opinion scale. Forms of the opinion-scale technique have been developed by Thurstone and Remmers. No scales have been published for appraising attitudes toward music, but teachers can adopt opinion scales and interest inventories for this purpose."¹² (p.406) This paragraph typifies the lack of specificity in discussions regarding affective domain assessment.
8. Teaching Music In The Secondary Schools (Charles Hoffer):¹³ The sixth chapter, "Evaluating the Results of Instruction" spans nineteen pages (pp. 105-124), with the topic "Affective Evaluation" receiving just two (pp. 117-118).
9. Teaching Music In Today's Secondary Schools (M.E.Bessom, A.M.Tatarunis, and S.L.Forcucci):¹⁴ The thirteenth chapter, "Measurement And Evaluation" takes twenty-six pages (pp. 292-318), however the topic of affective assessment is discussed in only three (pp. 309-312).
10. Creative Teaching of Music In The Elementary School (Dorothy Hickok and James A. Smith):¹⁵ There is no specific chapter assigned to assessment or evaluation. Reference to the topic is found under the subheading "Creative teaching is success- rather than failure-oriented". However the authors advocate a 'hold

¹¹ C. Leonhard and R. House, Foundations and Principles of Music Education, 2nd ed., (1972).

¹² Ibid, p. 406.

¹³ C. Hoffer, Teaching Music in the Secondary Schools, 2nd ed., (1973).

¹⁴ M. Bessom, A. Tatarunis, and S. Forcucci, Teaching Music in Today's Secondary Schools, (1974).

¹⁵ D. Hickok and J. Smith, Creative Teaching of Music in the Elementary School, (1974).

back' approach: "Even at the onset of the creative act, excessive evaluation may be construed by children as disapproval and may check the creative flow. The work of many researchers [this statement is not supported by references] would propose that evaluation and criticism of ideas be postponed until all ideas are out. This is often called the principle of 'deferred judgement'."¹⁶ Under a further subheading, "Music Evaluation", the authors suggest that "the best evaluation of any musical program is not only how well each child can sing or play, but does each child approach the music experience with anticipation and joy; does he react emotionally and intellectually to music learning activities; and does the music experience contribute to his creative development?"¹⁷ No details are supplied as to how the above evaluations might be undertaken.

11. Discovering Music: Developing the Music Curriculum in Secondary Schools (Keith Swanwick and Dorothy Taylor): ¹⁸ There is no specific section referring to evaluation or assessment. The affective objective 'valuing' is identified in the text, but only as an *outcome* of the education process. The main emphasis in this book is the understanding of musical concepts, and skill acquisition. The authors state: "While acknowledging that it is one of the main purposes of education to enhance and extend the range of what is valued, we also recognise that there can be no attempt to tackle value changes head-on in the classroom. Since we look for development here but cannot predict it, we can leave Values out of curriculum details while remembering that if music is not valued more because of our work we shall have been wasting our time."¹⁹ So they are advocating 'valuing' as being a major objective for music education, yet they feel that it may not be monitored or evaluated!

¹⁶ Ibid, p. 23.

¹⁷ Ibid, p. 42.

¹⁸ K. Swanwick and D. Taylor, Discovering Music: Developing the Music Curriculum in Secondary Schools, (1982).

¹⁹ Ibid, p.128.

12. Music in the Secondary School Curriculum (John Paynter):²⁰ This text is principally a report of the University of York's project of the same name. One of the key issues addressed is "Examinations and Assessment" (Chapter 6), but here Paynter is mainly concerned with incorporating an 'aural' element into assessment. The limitations of current examination practice is considered, but affective domain concerns are not highlighted.
13. Education and Music (Peter Fletcher):²¹ Although primarily following a philosophical/historical approach to the topic of music education, Fletcher does bring his ideas together in a more 'practical' vein in Chapter 18, "Priorities in Relation to Resources." Here, he suggests and discusses 'Aims', 'Objectives', 'Methods' etc of school music education. Despite this detail, however, affective aims (at best) are implied rather than stated, and assessment issues (both cognitive and affective) are not addressed.

The resulting problem, as far as Music Education is concerned, is a wide discrepancy between stated affective objectives for a particular course, and any subsequent evaluation of that course.

It is well documented that affective elements such as feelings, attitudes, and values are more difficult to test and assess than their cognitive counterparts; the result being that they tend to be credited with less and less importance. [This is often not the intentions of the educators, but the outcome is the same as when affective objectives are consciously disregarded.] Krathwohl, Bloom, and Masia describe this 'assessment difficulty' factor as

²⁰ J. Paynter, Music in the Secondary School Curriculum, (1982).

²¹ P. Fletcher, Education and Music, (1987).

contributing to an “erosion of affective objectives.” In their research, it was found that the original statements of various college courses gave “as much emphasis...to affective objectives as to cognitive objectives.”²²

*However, they report, as we followed some of these courses over a period of ten to twenty years, we found a rather rapid dropping of the affective objectives from the statements about the course and an almost complete disappearance of efforts at appraisal of student growth in this domain. It was evident to us that there is a characteristic type of erosion in which the original intent of a course or educational program becomes worn down to that which can be explicitly evaluated for grading purposes.*²³...

What then are some of the main difficulties and problems associated with assessment in the affective domain? A comprehensive list of headings assembled for the present study runs as follows:

- (1) A lack of constructive research.
- (2) The element of subjectivity in affective domain assessment, including the difficulty of finding objective descriptors and/or assessment tools.
- (3) The requirement for assessors to adopt flexible standards.
- (4) The problems associated with the ‘credibility gap’.
- (5) The issue of test anxiety.
- (6) The anonymity issue in affective domain assessment.
- (7) A lack of clearcut definitions with regard to statements of objectives, together with confusion arising from the multiplicity of terms employed in the literature.
- (8) The expectations of the students themselves, as far as affective assessment is concerned.
- (9) The ‘personal’ and private nature of the affective domain.

²² D. Krathwohl, B. Bloom, and B. Masia, Taxonomy of Educational Objectives, The Classification of Educational Goals - Handbook II: Affective Domain, (1964), p. 16.

²³ Ibid.

- (10) The cognitive versus affective domains, and associated problems of categorisation.

These issues will now be defined more closely:

(1) A lack of constructive research:

One of the disadvantages of affective domain assessment, as has been indicated in preceding discussion, is the lack of constructive research. Cognitive/objective-based assessment procedures enjoy an established, well-researched position in educational theory, and the sheer weight of these tends to dominate the rather few instances of complementary research undertaken in the affective domain [as evidenced by the literature].

The main effect of this imbalance is that the affective domain research risks (1) being taken to be subsidiary, and (2) falling into relative obscurity.

(2) The element of subjectivity in affective domain assessment:

Many educators, and indeed many established educational programmes, are not flexible enough to accommodate the notion of subjectivity playing an important role in the official assessment process. As observed by Whybrew,

...so many aspects of music, and consequently music education, are so difficult to approach through objective means. ... In the face of these difficulties, it may be that many music teachers are too ready to admit defeat in attempts to develop more systematic grading procedures. ²⁴

²⁴ W. Whybrew, Measurement and Evaluation in Music, (1962), p. 152.

(3) The requirement for assessors to adopt flexible standards:

Teaching for affective growth requires open-ended educational objectives, and does not contain the security of 'right and wrong' answers. Associated with this is the difficulty of applying standards, as 'global' standards are often inappropriate in the affective domain.

Krathwohl *et al* state:

*While there may be only one 'right' kind of achievement for an objective in the cognitive domain, there may be many 'right' behaviors equally correct in achieving an objective in the affective domain. With some objectives the 'right' answer can be judged only in terms of the criteria which the student sets for himself.*²⁵

The necessity for student-dictated criteria when teaching for affective growth is endorsed by Leonhard and House, as follows:

*Music education requires flexible standards adaptable to individual differences in ability and goals. Although the maintenance of standards represents a legitimate function of evaluation in music education, music educators should exercise great care in applying standards. When applied without regard to individual differences, standards may actually hinder musical development.*²⁶

²⁵ Krathwohl, Bloom and Masia, (1964), op.cit., p. 61.

²⁶ Leonhard and House, (1972), op. cit., p. 337.

(4) The problems associated with the ‘credibility gap’:

A term adopted by Eiss and Harbeck²⁷, the ‘credibility gap’ refers to the lack of certainty - particularly in the affective domain - “...that a given behavior is evidence of achievement of the desired objective...”²⁸. Bloom, Madaus and Hastings take up discussion in this area:

Perhaps the single most troublesome problem is that students’ actions or products specified as indicators of such affective constructs as interests, attitudes, and values might in fact be reflecting no more than a desire to please or to give the socially acceptable response. ²⁹

Hopkins and Stanley identify this problem as the ‘fakability’ factor. They suggest:

*Ability tests are keyed a priori; it is not possible to fake a high (or good) score (although of course one can fake a low score, that is, deliberately score poorly on the test). On affective measures one can usually fake in either direction...*³⁰

Interestingly, if subjects are moved to ‘fake’ their responses in order to appear positive towards music, then they are indicating (at the very least) that they rate highly being thought of as possessing that positive attitude. As Thurstone suggests:

We shall assume that it is of interest to know what people say that they believe even if their conduct turns out to be inconsistent with their professed opinions. Even if they are intentionally distorting their attitudes, we are

²⁷ A. Eiss and M. Harbeck, Behavioral Objectives in the Affective Domain, (1969).

²⁸ Ibid, p. 4.

²⁹ B. Bloom, G. Madaus, and J. Hastings, Evaluation To Improve Learning, (1981), p. 309.

³⁰ K. Hopkins and J. Stanley, Educational and Psychological Measurement and Evaluation, (1981), p.310.

*measuring at least the attitude which they are trying to make people believe that they have.*³¹

(5) The issue of test anxiety:

Traditional assessment procedures are frequently linked to 'testing' - and its acknowledged accompanying stresses and anxieties. While this may be of concern when assessing in the cognitive domain, it is in fact a major drawback in the affective domain. For in the case of the latter, undue psychological pressure on students is likely to cause more instability in results - as the students have more control over the direction of their responses.

(6) The anonymity issue in affective domain assessment:

Linked to the dilemma of the 'credibility gap' is the issue of anonymity versus subject identification in affective measurement. To elaborate, one of the major concerns of assessment in the affective domain is, for example, whether students' responses to a measurement instrument are influenced by the fact that they are required to identify themselves; as opposed to filing their returns anonymously. The problem of the 'credibility gap' is directly applicable here, as obviously the fact that students may manipulate their answers to suit their own purposes undoubtedly weakens the reliability of the affective assessment procedure.

It would be of value now to examine elements of this anonymity issue, and attempt to pinpoint areas of greatest weakness - particularly as far as formal affective assessment requirements are concerned.

For the purposes of this discussion, formal assessment as a procedure may be divided into two areas: 'means' and 'outcomes'. The 'means', or the actual assessment process, has a

³¹ L. Thurstone, The Measurement of Values, (1959), p. 217.

distinct set of problems relating to the anonymity issue that are not necessarily influenced by those associated with the 'outcomes' of the assessment.

One problem occurring with assessment in the affective domain - and this is not restricted to the administration of a formal measurement instrument - is often the inappropriateness of student 'prior knowledge' of assessment objectives:

...announcing the objectives in behavioral terms in advance will not work in the affective domain. Because of the individual's operant conditioning, he will show almost any desired behavior in order to win the approval of the teacher and/or to get good grades.³²

In other words, 'prior knowledge' of assessment objectives will strongly influence and direct student responses, thus distorting results.

This is in contrast to the effectiveness of giving students a set of desired behavioral objectives in the cognitive domain. Once the student is aware of the specific behaviors desired, he will usually strive to attain them.³³

And in doing so, the student will submit responses which are, presumably, indicative of cognitive level:

...it is assumed that a student who responds in the desirable way on a cognitive measure does indeed possess a competence which is being sampled.³⁴

Interestingly, a procedure recognised in one situation as being an aid to learning proves, in another, to be a hindrance.

³² Eiss and Harbeck, (1969), op. cit., p. 13.

³³ Ibid.

³⁴ Krathwohl, Bloom and Masia, (1964), op. cit., p. 17.

As far as the present discussion is concerned, the problem outlined above arises from the students reacting to the simple fact that they are being assessed, thereby presenting atypical responses or behaviours. In such instances, anonymity would minimise this overreaction, as the onus would not be directly on the individual to perform well in relation to others. However, herein lies the disadvantage. Anonymity may increase reliability, but at the same time, individual monitoring is lost.

In examining the second problem area associated with assessment in the affective domain - that being the 'outcomes' of the assessment process - it is apparent that the anonymity issue is even more significant. Invariably, students are patently aware of the ultimate object of any assessment procedure, and in many cases are more concerned with achieving the right grade, or making the right impression, than in revealing something personal which might supposedly 'incriminate' them. Their aim may be to impress a teacher or friends, or to hide an opinion which it is feared may be socially unacceptable, or to achieve a 'standard' which leads to a higher good.

(7) A lack of clearcut definitions with regard to objectives statements, together with confusion arising from the multiplicity of terms employed in the literature:

Research into assessment in the affective domain is very much hampered by a multiplicity of terms - many of which are either lacking in definition, or abounding in ambiguities. Many terms, for example 'interests', 'values', 'attitudes', and 'appreciations', are used interchangeably, without much consistency being reached with past writings. The effects of this problem, predictably, is added confusion for educators and later researchers. Thus the apparent intangibility of the affective domain is perpetuated further.

To compare this situation with that of research in the cognitive domain highlights startling deficiencies. Krathwohl *et al* outline these:

More than two decades of work on cognitive objectives have produced specific and meaningful results. Few serious workers now use such terms as “critical thinking,” “problem solving,” or “higher mental processes,” as statements of objectives. These terms may be used to describe large goals and aims of education, but in describing the objectives of a course with specific sequences of learning experiences, curriculum makers have more recently made use of terms like “application of principles,” “interpretation of data,” “skill in recognising assumptions,” etc.³⁵

Conversely, the authors report, in the affective domain:

...the situation... is so primitive that little in the way of meaning is at present conveyed by statements of objectives.³⁶

By way of example, an objective statement from the affective domain is cited, viz : “The student should become interested in good books.” They continue:

What is meant by “interest” may range from simply knowing that the object exists to a passionate devotion to this type of object or activity.³⁷

Admittedly the source currently under discussion was published back in 1964, nevertheless it is significant to note its continuing relevance after a span of nearly thirty years. The following objective statement, for example, appears in a 1987 music educational policy document:

³⁵ Ibid, p. 21.

³⁶ Ibid.

³⁷ Ibid, p. 22.

*All students should be encouraged to take an interest in, and be open minded about, the music of our own time.*³⁸

Words such as 'encouraged', 'interest', and 'openminded' are somewhat vague terms without specific application. Hence the broad, general nature of such an objective does little for the educator requiring guidance and support in teaching for and assessing affective growth.

(8) The expectations of the students themselves, as far as affective assessment is concerned:

Traditionally, the emphasis in education has been on cognitive testing, not affective assessment. Students expect to have their knowledge and their comprehension skills examined, but not their attitudes, feelings, interests and values. They know that their course grades and examination percentages comprise the major, if not only, component of their formal academic record. As suggested by Marion Metcalfe, in her article "Towards the Condition of Music":

*As long as the examination system exists in our society as the accepted way of validating educational experience it cannot be ignored or sidestepped by those who wish to promote their 'subject'. The Gulbenkian Report of 1982 was the most powerfully and weightily presented argument for the arts in education yet to appear in the public arena of debate, but to pupils, parents and employers argument alone is not enough; it is the magic pieces of paper labelled 'certificate' that count.*³⁹

³⁸ Draft Syllabus for Music Education, Early Childhood to Form Seven, New Zealand Department of Education, (1987), p. 13.

³⁹ M. Metcalfe, "Towards the Condition of Music: The Emergent Aesthetic of Music Education" in Living Powers: The Arts In Education, ed: P. Abbs, (1987), pp 112.113.

With regard to the New Zealand education system, the liberal use of formal examinations to monitor student development is undeniably well established. David Sell refers to both the school music setting, and the field of private music-teaching as harbouring a “British-based pre-occupation with examinations.”⁴⁰ He continues:

With regard to examinations in music, New Zealand children are probably the most examined in the world. ⁴¹

In sum, for most New Zealand students the ‘passing’ of official examinations is accepted as being a primary goal of education. Consequently, the expectations of the students, and the education system as a whole, have traditionally combined to undermine the status of formal affective domain assessment.

(9) The ‘personal’ and private nature of the affective domain:

Affective assessment seeks to reveal aspects of the ‘inner’ person; an exercise which is not automatically welcomed by adolescents. Linked with this is perhaps an even more important issue, as raised by Krathwohl *et al.*:

A much more serious reason for the hesitation in the use of affective measures for grading purposes comes from somewhat deeper philosophical and cultural values. Achievement, competence, productivity, etc., are regarded as public matters. Honors are awarded for high achievement, honor lists may be published by the Dean, and lists of National Merit Scholarship winners may be printed in newspapers. In contrast, one's beliefs, attitudes, values and personality characteristics are more likely to be regarded as private matters.... Each man's home is his castle, and his interests, values, beliefs, and personality may not be scrutinized unless he

⁴⁰ D. Sell, “Five Englishmen - The British Influence on Music Education in New Zealand Between the Wars” in *Studies in Music Education*, ed: D. Sell, (1986), p. 15.

⁴¹ Ibid, p. 16.

voluntarily gives permission to be revealed. This public-private status of cognitive vs. affective behaviors is deeply rooted in the Judaeo-Christian religion and is a value highly cherished in the democratic traditions of the Western world. ⁴²

(10) The cognitive versus affective domains, and associated problems of categorisation:

In order to discuss the second part of this issue, it is necessary to further define the terms 'cognitive' and 'affective' as employed in the context of the current study.

The affective domain as a concept is rather more broad than that embraced by the term 'affective response' as used, for example, by Lundin in An Objective Psychology of Music.⁴³ Lundin links as synonymous the terms 'affective' and 'feeling' when describing responses to music, and it is here that a specific distinction can be made. In Lundin's words:

An 'affective' response is one in which the stimulus has made some definite change in the organism. It may be active and energy-consuming, but it has no effect on the stimulus object. ⁴⁴

Affective response to music, then, has an episodic immediacy about it that is not necessarily inferred by the concept of the affective domain and music. One might respond affectively to a particular experience or stimulus, - the result being a temporary mood, emotion, or feeling alteration. However, this may not necessarily have any bearing on one's overall interests or values. Affective responses, in fact, constitute only a small part of the affective domain. [Recall the definition by Ringness, on p.1]

⁴² Krathwohl, Bloom and Masia, (1964), op. cit., pp. 17-18.

⁴³ R. Lundin, An Objective Psychology of Music, (1967).

⁴⁴ Ibid, p. 151.

A point that is now emerging, with regard to assessment procedures in the affective domain, is that areas of potential educational interest are vast. However, too often issues and objectives are restricted to such as the following: whether the student ‘likes or dislikes’ music; or how the student responds affectively to specific musical stimuli (such as described by Lundin, above).

The problem referred to of terminology confusion and primitiveness was addressed by Krathwohl, Bloom and Masia who, in 1964, undertook to comprehensively and hierarchically classify objectives in the affective domain, along the lines of the work already completed in the cognitive domain. [Although a fuller discussion of the resulting Taxonomy will occur in Chapter II of this study; an indication of its significance is necessary at this point.] In this project⁴⁵, the nature of the affective domain was outlined, with the contributing elements and their interrelationships defined with constructive clarity. The aim was to identify a ‘continuum’ of affective objectives “...that would provide a means of ordering and relating the different kinds of affective behavior.”⁴⁶

Krathwohl *et al* agreed upon three ‘divisions’ of educational objectives, respectively ‘cognitive’, ‘affective’, and ‘psychomotor’. These divisions, or domains, are defined in full as follows:

- (a) **Cognitive:** Objectives which emphasize remembering or reproducing something which has presumably been learned, as well as objectives which involve the solving of some intellectual task for which the individual has to determine the essential problem and then reorder given material or combine it with ideas, methods, or procedures previously learned. Cognitive objectives vary from simple recall of

⁴⁵ Krathwohl, Bloom and Masia, (1964), op. cit.

⁴⁶ Ibid, p. 24.

material learned to highly original and creative ways of combining and synthesizing new ideas and materials.

- (b) **Affective:** Objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomena to complex but internally consistent qualities of character and conscience. [We found] a large number of such objectives in the literature expressed as interests, attitudes, appreciations, values, and emotional sets or biases.
- (c) **Psychomotor:** Objectives which emphasize some muscular or motor skill, some manipulation of material and objects, or some act which requires a neuromuscular co-ordination.⁴⁷

These domains as defined appear to indicate clearcut divisions among groups of objectives, yet Krathwohl *et al* are the first to point out the deficiencies of such a system. The main problem, of course, is that one cannot simply 'divide' human processes into separate categories - overlapping is an integral factor. It was found that :

...although one could place an objective very readily in one of the three domains or classes, no objective was entirely devoid of some components of the other two classes. ⁴⁸

For example, the performer who gains immense pleasure [affective response] from playing a musical instrument will also be utilising both music reading skills [cognitive ability] and physical skills [psychomotor domain]. The question the authors raise is "...whether a human being ever does thinking without feeling, acting without thinking, etc..." They

⁴⁷ Ibid, pp. 6-7.

⁴⁸ Ibid, p. 8.

continue : “It seems very clear that each person responds as a ‘total organism’ or ‘whole being’ whenever he does respond.”⁴⁹

By the same token, it could be argued that if cognitive objectives contain an affective component, and vice versa, then there is no need to subscribe to separate categories. However, as is pointed out by Krathwohl *et al*:

The affective domain is useful [and, one might suggest, necessary] in emphasizing the fact that affective components exist and in analyzing their nature. ⁵⁰

It is interesting at this point to consider the Aristotelian theory of ‘Eudaimonia’ (Happiness), and realise that the identification of affective components in cognitive objectives is by no means peculiar to modern-day educational thinking. Indeed, it was Aristotle’s view that man’s highest attainment is an affective goal, and one to which we all ultimately aspire. The following extract is from ‘The Nicomachean Ethics’ :

As there is evidently a plurality of ends, and as some of these are chosen only as means to ulterior ends..., it is clear that not all ends are final. But the supreme good must of course be final. Accordingly, if there is only one final end, this will be the good that we are seeking; and if there is more than one such end, the most complete and final of them will be this good. Now we call what is pursued as an end in itself more final than what is pursued as a means to something else; and what is never chosen as a means we call more final than what is chosen both as an end in itself and as a means; in fact, when a thing is chosen always as an end in itself and never as a means we call it absolutely final. Happiness (eudaimonia) seems, more than anything else, to answer to this description: for it is something we choose always for its own sake and never for the sake of something else; while honor,

⁴⁹ Ibid, p. 7.

⁵⁰ Ibid, p. 49.

pleasure, reason, and all the virtues, though chosen partly for themselves (for we might choose any one of them without heeding the result), are chosen also for the sake of the happiness which we suppose they will bring us. Happiness, on the other hand, is never chosen for the sake of any of these, nor indeed as a means to anything else at all. ⁵¹

The philosophy of Aristotle recounted above relates directly to a statement by Krathwohl *et al* regarding modern day educational practice:

*Even though the whole school system rewards the student more on a 'can do' than on a 'does do' basis, it is the latter which every instructor seeks.*⁵²

To explain further, a traditional, cognitive-based assessment system would take into account the level of ability demonstrated by a student in a particular activity: in other words, what they 'can' do, or are capable of. What is not recorded, or given credit for, is the fact that the student may be highly motivated to seek out the particular activity - despite being relatively limited in ability. As an example, the person who demonstrates a 'natural' talent for musical performance, and who is fortunate enough to progress with little or no effort, is deemed a more 'successful' student than one who practises diligently but shows no real talent. Hence the 'can do' student is rewarded over the 'does do' student, but ironically, as educators we are, informally, extremely interested in whether our students enjoy a particular activity enough to pursue it of their own volition.

In short, one of the main concerns regarding the relationship of the affective and cognitive domains is the issue of 'means' and 'ends'. It can be perfectly acceptable to employ affective means to achieve cognitive ends; so too can it be appropriate to utilise cognitive means to achieve affective ends: the problem occurs however when the cognitive elements

⁵¹ P. Wheelwright, *Aristotle*, (1951), p. 167.

⁵² Krathwohl, Bloom and Masia, (1964), op. cit., p. 60.

assume an all-consuming importance, and the affective 'ends' are relegated to mere byproducts of the educational process.⁵³

As stated by Tyler then:

*Human feelings are important both as means and ends in education.*⁵⁴

It would be very easy to become disillusioned about developing new directions for the assessment of affective elements in music education. The danger occurs when educators take the problems outlined above to be insurmountable and conclude that affective assessment per se has no place in the formal student assessment structure.

Edwards and Edwards report that "the consideration of attitudes as an aspect of pupil achievement has long been recognised but all too frequently omitted from formal pupil evaluation instruments."⁵⁵ Others are happy to make group affective assessments so as to evaluate aspects of their teaching programmes (Hopkins and Stanley, for example, conclude that "Anonymity is an important ingredient in valid attitude assessment in any situation in which the subject may be rewarded for, or embarrassed by, certain responses."⁵⁶); but very rarely are structured individual assessments undertaken with respect to affective development.

The view of the present study with regard to, for example, the anonymity issue, is that special efforts to assess individual affective growth must be made. It is suggested that the solution may begin with seeking to minimise the negative influences at work on the psyches of the students. According to Hopkins and Stanley, "Most affective measures

⁵³ Ibid, p. 57.

⁵⁴ R. Tyler, "Assessing Educational Achievement in the Affective Domain", Measurement in Education, (1973), p. 2.

⁵⁵ J. Edwards and M. Edwards, "A Scale to Measure Attitudes Toward Music", Journal of Research in Music Education, (1971), p. 230.

⁵⁶ Hopkins and Stanley, (1981), op. cit., p. 310.

require self reports that are fakable; thus if the assessment is to be valid, it must be obtained in such a way that there is no incentive to be untruthful.”⁵⁷ Eiss and Harbeck discuss placing the student in situations “... where he is not under any obligation to give the desired response.”⁵⁸

If one were to minimise negative influences of affective domain assessment, one would need to begin by focussing on the following:

- (1) formative as distinct from summative evaluation, so as to emphasise the ‘feedback’ role of assessment rather than the ‘judgemental’ one,
- (2) procedures which are not concerned with yoking a student’s growth to the achievements and levels of others,
- (3) avoiding the conventional ‘testing’ situation, so helping to ease the stress of traditional assessment procedures.

2. IMPORTANCE OF THE TOPIC

We return now to the broader issue of music education and affective domain assessment, and attempt to argue the case for further research by posing some important questions. But first, in order to obtain a balanced inquiry, one needs to be aware of the following cautionary philosophy. In responding to the contemporary call for accountability in education, it would be very easy to develop a single minded approach to assessing the myriad of elements that contribute to a valid curriculum. The danger here is that one could lose sight of the diversity in education, by adopting a blanket policy that betrays the inherent nature of some components. One is reminded of the oft-quoted Thorndike statement:

⁵⁷ Ibid.

⁵⁸ Eiss and Harbeck, (1969), op. cit., pp. 13-14.

Whatever exists exists in some amount and can be measured."⁵⁹

By following such a maxim, we risk viewing the human being as an 'input- output' mechanism - a term adopted by Reimer when pointing out the weaknesses of Behaviorism and music education.⁶⁰ Reimer's discussion highlights the dangers of adopting a 'blanket policy' approach:

*Behaviorism is most concerned with a specifiable, measurable, overt, dependent response to a chosen stimulus. Such a phenomenon, by definition, is a lower-order behavior. It is manifested clearly by machines and almost as clearly by lower-order organisms such as rats and cats. ... Since a stimulus-response unit at this level is the common coin of the realm in Behaviorism, a strong tendency exists to make the process of education one of a steady accretion of stimulus-responses.*⁶¹

As far as the present discussion is concerned, merely by seeking to redress the intangibility of the affective domain we may be jeopardizing its very essence. For, primarily, assessment objectifies human processes. It is a way by which one may 'be viewed', by means of an external procedure. As defined by Rowntree:

*...assessment in education can be thought of as occurring whenever one person, in some kind of interaction, direct or indirect, with another, is conscious of obtaining and interpreting information about the knowledge and understanding, or abilities and attitudes of that other person.*⁶²

Nevertheless, this does not mean that the assessment measures themselves must be exclusively objective. The Encyclopedia of Educational Evaluation states that 'assessment'

⁵⁹ In: Leonhard and House, (1972), op. cit., p 390.

⁶⁰ B. Reimer, Toward an Aesthetic Education, (1971), p. 67.

⁶¹ Ibid, pp. 69-70.

⁶² D. Rowntree, Assessing Students, How Shall We Know Them?, (1987), p. 4.

is 'multitrait - multimethod'; in other words, "...it focuses on a number of variables judged to be important and utilises a number of techniques to assay them."⁶³ On the issue of subjectivity in assessment Radocy suggests that:

*Music educators should recognize that all measurement procedures they are likely to employ become subjective at some level, in construction, application, and/or interpretation. Many important aspects of performance and teaching cannot be measured by counting units; human judgements are necessary. ... Nevertheless, by employing certain psychophysically - based procedures it is possible to quantify that which is not readily countable if one can accept matching impressions across different sensory modalities as a basis for measurement.*⁶⁴

The use of the word 'measurement' by Radocy illustrates the semantic difficulties that besiege affective domain writing, for in fact there is some question as to whether 'measurement' is an appropriate term for arts education. According to Kleinig, 'measurement' covers any arithmetic or symbolic means of representing what has been learned. He then follows with a suggestion that perhaps "learning and other human capacities" are not measurable, in the mathematical sense of the word.⁶⁵ Illich endorses this concern when he writes:

The institutional values school instils are quantified ones. School initiates young people into a world where everything can be measured, including their imaginations, and, indeed, man himself. But personal growth is not a measurable entity. It is growth in disciplined dissidence, which cannot be

⁶³ S. Anderson, S. Ball and R. Murphy, The Encyclopedia of Educational Evaluation, (1975), pp. 26-27.

⁶⁴ R. Radocy, "On Quantifying The Uncountable in Musical Behaviour", Bulletin of the Council for Research in Music Education, (1986), p. 30.

⁶⁵ J. Kleinig, Philosophical Issues in Education, (1982), p. 178.

measured against any rod, or any curriculum, nor compared to someone else's achievement. ⁶⁶

Despite such concerns, however, there are certain questions to be asked which, when answered, provide considerable justification for further research into affective domain assessment in Music Education.

To begin with, why focus at all on affective elements in assessment?

Essentially, it would be very rare indeed to discover a music education course that was not concerned with student affective growth and development, and almost as rare for affective objectives to be omitted from the formal course outline. It follows therefore that if affective elements, which are integral to music education, are not focussed on in the assessment process then their inclusion in a course outline is an empty gesture. As suggested by Ebel:

*The cultivation of variously presumably desirable attitudes is frequently mentioned as an educational objective for a course or a school. Less often does the teacher or the school have a systematic program for the attainment of such objectives. Almost never is a serious attempt made to measure the extent to which such objectives have been achieved.*⁶⁷

*The evaluation of appreciation and attitudes is essential to successful music teaching. It not only enables the teacher to appraise results in important areas of musical learning but also provides him with valuable insights into the nature and needs of his pupils, both as a group and as individuals.*⁶⁸

⁶⁶ I. Illich, Deschooling Society, (1971), p. 40.

⁶⁷ R. Ebel, Essentials of Educational Measurement, (1972), p. 527.

⁶⁸ Leonhard and House, (1972), op. cit., p. 351.

Colwell draws attention to a point that one should remain aware of when developing objectives for evaluation in the affective domain:

*A [second] problem is that evaluation may not be able to differentiate between value responses that come as a result of maturity and those that are the consequence of education. The concerns of society that are satisfied by merely living in society for twenty years should be recognized, so that the school does not fritter away its resources upon goals that are achieved naturally.*⁶⁹

This is a valid caution; however, educators need to ensure - through the assessment process - that affective objectives are met, and are not simply allowed to subsist in the shadow of cognitive domain emphases. In too many situations this imbalance exists, fuelled by educators who may well consider affective elements to be of value, but who relegate them to the status of 'happy byproducts' of the educational system. Taylor is one writer who rejects such a relegation, suggesting instead a somewhat contrasting approach:

*Given time, any competent music faculty can discover talent or its absence. ... It is the personality factors that are so often unknown and unpredictable. A good natural voice, a sensitive ear, and a feeling for musical values come to light with little difficulty, but such matters as integrity, energy, and persistence are also determinants of success or failure, no matter how great - or small - the talent. A complete psychological prognosis must account for them.*⁷⁰

⁶⁹ R. Colwell, The Evaluation of Music Teaching and Learning, (1970), p. 130.

⁷⁰ C. Taylor, "Techniques for the Evaluation of Musical Status", Journal of Research in Music Education, (1963), p. 62.

A second question to be concerned with is why test specifically for attitude levels?

To be more particular, why must we isolate aspects of the affective domain for assessment when they may well be catered for in traditional cognitive-based assessment procedures? It is often believed, for example, that high achievement in a course is linked to positive attitudes and values. Krathwohl *et al* report that "...there still persists an implicit belief that if cognitive objectives are developed, there will be a corresponding development of appropriate affective behaviors."⁷¹ Edwin Gordon states: "From the little evidence we have, we can only assume that one's interest in music does affect his success in music."⁷² However this is not necessarily always the case. Birney, for example, in his article, "The Effects of Grades on Students", suggests: "In courses with low interest, high grades have the effect of lessening the amount of study. In courses with high interest, high grades may stimulate greater interest."⁷³ The point to be made here is that any two instances of high achievement in a course may well be corresponding with widely varying levels of interest. In other words, cognitive levels are not indicative of affective levels. Krathwohl *et al* support this view when they state that "...under some conditions the development of cognitive behaviors may actually destroy certain desired affective behaviors and ... instead of a positive relation between growth in cognitive and affective behavior, it is conceivable that there may be an inverse relation between growth in the two domains."⁷⁴

In his research, Bullock found that "...several studies...demonstrated that attitudinal change is unrelated to growth in knowledge of course content", and cites a number of cases to support this.⁷⁵ Similarly Asmus [1980], in his investigation into the premise "...that student affect toward a course of instruction will dictate, in part, cognitive performance in

⁷¹ Krathwohl, Bloom and Masia, (1964), op. cit., p. 20.

⁷² E. Gordon, The Psychology of Music Teaching, (1971), p. 33.

⁷³ R. Birney, "The Effects of Grades on Students", Journal of Higher Education, (1964), p. 96.

⁷⁴ Krathwohl, Bloom and Masia, (1964), op. cit., p. 20.

⁷⁵ W. Bullock, Construction and Evaluation of a Test of Musico-Aesthetic Attitude, (1971), p. 5.

the course”⁷⁶, discovered that the opposite was in fact the case: that “...affect did not appear to predict performance on cognitive variables.”⁷⁷ On discovering that “Affect ... does not appear to be crucial in promoting cognitive change”⁷⁸, Asmus takes his concluding remarks to perhaps unwarranted lengths:

*The results of this investigation suggest that music educators could best promote learning by concentrating their efforts on music content rather than on trying to influence student attitudes and opinions toward the course of music instruction. The time spent in influencing student feelings and perceptions may be used more wisely by focussing on skills necessary for music literacy.*⁷⁹

Asmus does not appear to acknowledge that some ‘courses of music instruction’ focus in part on affective development: which, as his investigation showed, is not automatically synonymous with cognitive growth.

Thorndike and Hagen caution:

*It is important not to confuse measures of interest and ability. Interest measures tell us nothing directly about ability and, generally speaking, the relationship between interests and abilities are quite low.*⁸⁰

Pogonowski discovered from her study into “Attitude Assessment of Upper Elementary Students in a Process-Oriented Music Curriculum” that “... classroom music attitudes and music aptitude (the Melody and Tempo subtests of the MAP) were essentially unrelated to

⁷⁶ E. Asmus, “Empirical Testing of an Affective Learning Paradigm”, Journal of Research in Music Education, (1980), p. 143.

⁷⁷ Ibid, p. 151.

⁷⁸ Ibid.

⁷⁹ Ibid.

⁸⁰ R. Thorndike and E. Hagen, Measurement and Evaluation in Psychology and Education, 4th ed., (1977), p. 416.

each other.”⁸¹ The question “Are there any relationships between low, average, or high musical aptitude and classroom music attitudes?”, was one of four in an investigation undertaken by Pogonowski to “determine whether attitudes among students engaged in a Process-Oriented Music Curriculum (POMC) differ as a function of gender, grade level, socioeconomic status, and musical aptitude.”⁸² “The three measures used to collect data were a) the Music Class Attitude Index [“modelled after the inventory used by Nolin” - see Chapter IV], b) the Music Attitude Inventory [“an instrument developed by the author for use in this study”], and c) Melody and Rhythm subtests of the Musical Aptitude Profile [Gordon].”⁸³ Studies such as the aforementioned support the view of there being little relationship between affective and cognitive growth, indicating a necessity for educators to conduct specifically-designed assessments in each domain.

A third question relates back to a discussion on p.26, viz: Why is there a need for structuring an approach to the assessment of affective elements?

The answer here is that too often assessment in the affective domain is neglected, due to its apparent intangibility. Hopkins and Stanley observed that “Although much attention is given to the assessment of cognitive objectives, rarely is any *systematic* effort directed toward the evaluation of affective objectives.”⁸⁴ And more specifically, according to Froelich, “Anyone who attempts to develop a tool for measuring affective learning enters a novel area within music education research.”⁸⁵

Attention was drawn earlier (p.6) to the ‘assessment difficulty factor’ which exists in the affective domain - a factor which prompted the classificatory work of Krathwohl *et al.*

⁸¹ L. Pogonowski, “Attitude Assessment of Upper Elementary Students in a Process-Oriented Music Curriculum”, Journal of Research in Music Education, (1985), p. 256.

⁸² Ibid, p. 247.

⁸³ Ibid, pp. 247, 249, 250.

⁸⁴ Hopkins and Stanley, (1981), op. cit., p. 289.

⁸⁵ H. Froelich, “Raymond Charles Wifler: The Development and Application of an Instrument to Identify Music-Affect Shift Resulting from a College Music Appreciation Course: A Review”. Bulletin of the Council for Research in Music Education, (1980), p. 49.

There is a need for continued research into all aspects of assessment in the affective domain, otherwise defeatist attitudes such as the following will remain, thereby hindering educational development:

*If achievement in certain aspects of a program is difficult to assess, it might be suggested that particular goals are in need of clarification or that they are completely unrealistic and therefore might be abandoned. It makes little sense to say that something can be taught which does not lend itself to evaluation.*⁸⁶

One must surely challenge this last statement with the question: is it the inappropriateness of the 'something' being taught, or the inadequacy of the 'evaluation' process which is the restricting factor? If an objective 'does not lend itself to evaluation', it may well be that the evaluation procedures available are insufficient and in need of updating, rather than that the objective itself is unrealistic.

⁸⁶ Gordon, (1971), op. cit, p. 133.

CHAPTER II - PURPOSE OF THE STUDY

The purpose of this study is twofold: the first purpose (as discussed thus far) has been to examine the issue of assessment in the affective domain, with particular reference to New Zealand secondary school general music education. In order to support this examination, two further areas of importance need to be presented. The first of these pertains specifically to the development of secondary school general music education in New Zealand: this discussion may be found in Chapter III. The second 'area of importance' involves a review of the literature and past research associated with issues of affective domain assessment and music. This review (which constitutes Chapter IV) is intended to set the scene for further - and more importantly - related work.

The second purpose of the study seeks to make a practical contribution to the assessment of elements in the affective domain, by addressing problems and concerns raised by the examination of the various affective domain issues. The 'practical contribution' will take the form of the development of a group music attitude assessment instrument, for use in New Zealand secondary school general music classes. This second 'purpose of the study' will now be introduced.

1 ASSESSMENT OF ELEMENTS IN THE AFFECTIVE DOMAIN - AN INTRODUCTION

Up to this point, the discussion has been shaped to portray the great need for sound, appropriate, and educationally valid methods of assessing elements in the affective domain. It has been argued that the cognitive domain has hitherto gained wider attention from researchers, to the detriment of the affective domain - both in teaching and in evaluation aspects. The difficulties and problems associated with assessment in the affective domain have been acknowledged; however it is the belief of the present researcher that these may not be insurmountable.

*Educational objectives in this [the affective] domain tend to be statements of desirable but undefined virtues. As long as the affective objectives remain in this empty and airy limbo, there is little that is likely to be done in the school either in evaluation or in the providing of appropriate learning experiences. If affective objectives can be defined with appropriate precision, we believe it may be no more difficult to produce changes in students in this domain than it has been in the cognitive domain.*⁸⁷

Of major interest to any study involving investigation of aspects of the affective domain is the work of Krathwohl, Bloom and Masia, as outlined in their book, Taxonomy of Educational Objectives, The Classification of Educational Goals. Handbook II: Affective Domain.⁸⁸ This Handbook was prepared in response to what was perceived by the authors as being a great need to formalise aspects of the affective domain - in terms of teaching and evaluation objectives.

*What is missing is a systematic effort to collect evidence of growth in affective objectives which is in any way parallel to the very great and systematic efforts to evaluate cognitive achievement.*⁸⁹

As mentioned earlier (p.18), the work of Krathwohl *et al* sought to 'comprehensively and hierarchically classify objectives in the affective domain', thereby providing "...a means of ordering and relating the different kinds of affective behavior."⁹⁰

The decision to identify three separate domains, those being cognitive, affective and psychomotor, was made in order to highlight the specific features of each. Furthermore:

There still persists an implicit belief that if cognitive objectives are developed, there will be a corresponding development of appropriate

⁸⁷ D. Krathwohl, B. Bloom and B. Masia, Taxonomy of Educational Objectives, The Classification of Educational Goals. Handbook II: Affective Domain, (1964), p. 76.

⁸⁸ D. Krathwohl, B. Bloom and B. Masia, (1964), op. cit.

⁸⁹ Ibid, p. 16.

⁹⁰ Ibid, p. 24.

*affective behaviors. ... The evidence suggests that affective behaviors develop when appropriate learning experiences are provided for students much the same as cognitive behaviors develop from appropriate learning experiences.*⁹¹

While it was acknowledged that the three divisions were somewhat arbitrary, the authors felt that they reflected the distinctions that were frequently made between “problem solving and attitudes, between thinking and feeling, and between acting and thinking and feeling.”⁹²

Following the research made into classifying objectives in the cognitive domain [as reported in Bloom *et al.*, Taxonomy of Educational Objectives: Cognitive Domain], Krathwohl *et al* were charged with equalling the exercise, this time focussing on the affective domain. The authors stated, in support of the task:

*If affective objectives and goals are to be realized, they must be defined clearly; learning experiences to help the student develop in the desired direction must be provided; and there must be some systematic method for appraising the extent to which students grow in the desired ways.*⁹³

The aim was to identify objectives specific to affective development, which would be capable of being arranged in the form of a continuum. Krathwohl *et al* reported:

*...the materials from which this continuum was to be deduced were the objectives dealing with interests, attitudes, values, appreciation, and adjustment.*⁹⁴

⁹¹ Ibid, p. 20.

⁹² Ibid, p. 7.

⁹³ Ibid, p. 23.

⁹⁴ Ibid, p. 24.

They continued by examining each of these in detail, then moved to a discussion introducing the concept of internalization. This last "...is viewed as a process through which there is at first an incomplete and tentative adoption of only the overt manifestations of the desired behavior and later a more complete adoption."⁹⁵ The process of 'Internalization' in fact becomes central to the development of the Taxonomy, as it forms the basis of the continuum, linking the objectives as they progress through a five - category hierarchy model.

The five categories (and their subcategories) of the Taxonomy are as follows:

- 1.0 **Receiving** (Attending)
 - 1.1 Awareness
 - 1.2 Willingness to receive
 - 1.3 Controlled or selected attention

- 2.0 **Responding**
 - 2.1 Acquiescence in responding
 - 2.2 Willingness to respond
 - 2.3 Satisfaction in response

- 3.0 **Valuing**
 - 3.1 Acceptance of a value
 - 3.2 Preference for a value
 - 3.3 Commitment (conviction)

⁹⁵ Ibid, p. 29.

4.0 **Organization**

4.1 Conceptualization of a value

4.2 Organization of a value system

5.0 **Characterization**

5.1 Generalized set

5.2 Characterization

[The subcategories, while effecting arbitrary divisions of the main taxonomic levels, do occur at points along the continuum, therefore are hierarchical in themselves.]

Krathwohl *et al* described the internalization process as it relates to their affective taxonomy as follows:

The process begins when the attention of the student is captured... As he pays attention ...[to that which interests]...he differentiates it from the others present in the perceptual field. With differentiation comes a seeking out of the phenomenon as he gradually attaches emotional significance to it and comes to value it. As the process unfolds he relates this phenomenon to other phenomena to which he responds that also have value. This responding is sufficiently frequent so that he comes to react regularly, almost automatically, to it and to other things like it. Finally the values are interrelated in a structure or view of the world, which he brings as a 'set' to new problems.⁹⁶

⁹⁶ Ibid, p. 33.

There are, in effect, several features, or levels, of the internalization continuum which may be identified [Krathwohl *et al* referred to the continuum as being 'multidimensional'⁹⁷], as the following paraphrase reports:

- (1) the simple to complex aspect
- (2) the concrete to abstract aspect
- (3) the external to internal control transition
- (4) the increasing emotional component
- (5) the conscious to unconscious aspects

These will be referred to in the following summary (of the five categories in ascending order of their degree of internalization) of the affective Taxonomy:

1.0 RECEIVING (Attending)

This first category refers to the very initial stages of affective growth; all that is required is that the subject is open to receiving certain stimuli or phenomena. It ranges from the subject merely being aware of those stimuli, to the point where they might actually direct their own attention to them. At this level, the subject feels no particular heightened feeling towards the phenomena (low emotional content), the control is still very much external (for example under the guidance of a teacher), and the conscious aspect is very strong - almost to the point of being a cognitive behavior.

2.0 RESPONDING

The 'Responding' category is concerned with behaviors which go beyond merely attending to stimuli or phenomena when directed; it continues on from the point of 'selected attention' and implies that the subject is experiencing some amount of self-

⁹⁷ Ibid, p. 31.

motivation. This motivation is recognised in level 2.0 to the extent that the subject feels satisfaction in responding to the phenomena (an increase in emotional content). One of the educational objectives at this point is that a student will seek out certain activities in response to that feeling of satisfaction. The commitment at this level, however, is still fairly low - the subject does not feel enough for the phenomena to regard it as a 'value'.

3.0 VALUING

'Valuing' refers to an increase in affective response to the extent that the subject 'ascribes worth' to a particular phenomenon or stimulus. The key words for this level are 'consistency', 'stability', and 'commitment', and the implication is that the subject will act "to further the thing valued in some way, to extend the possibility of his developing it," and "to deepen involvement with it and with other things representing it"⁹⁸. There are the beginnings of 'internalization', an increase in the complexity of the affective relationship, and many behaviors associated with 'valuing' are unconscious ones.

4.0 ORGANIZATION

At this level, the subject is beginning to perceive and examine held values in terms of a wider cosmos. Ideas and values are abstracted, and are brought "...into an ordered relationship with one another"⁹⁹, thereby constituting a 'value system' from which the subject may approach all other affective experiences.

5.0 CHARACTERIZATION

"The organization of values, beliefs, ideas and attitudes into an internally consistent system is called characterization. This goes beyond merely determining interrelationships among various values; it implies their organization into a total

⁹⁸ B. Bloom, G. Madaus and J. Hastings, Evaluation to Improve Learning, (1981), appendix.

⁹⁹ Ibid.

philosophy or world view.”¹⁰⁰ ‘Characterization’ represents the epitome of affective growth, and “can best be illustrated by the dedicated artist, scientist, statesman, teacher, ... whose life and decisions all revolve around a single focus.”¹⁰¹ The progressions from simple to complex, from concrete to abstract, from external to internal control, and from conscious to unconscious are now complete.

As far as the placement of ‘attitude’ on the Taxonomy continuum is concerned, one may turn to a chart, compiled by Krathwohl *et al* ¹⁰², which identifies and relates terms commonly used when describing affective domain development. It illustrates, using arrows, exactly where the different terms of interest, appreciation, attitude, value, and adjustment [see below for definitions] fall between Level One (Receiving), and Level Five (Characterization). Attitude, as does Value, spans Levels 2.2 through to Level 4.1, and overlaps the top ends of Appreciation and Interest, and the lower stages of Adjustment. The authors comment:

All the terms overlap one another in meaning in the middle range of the Taxonomy continuum. No specificity can be gained by replacing one term by another in this range, and the possibilities for confusion are great. ... every term is overlapped by at least one other term for a major part of that part of the continuum it describes. ¹⁰³

Earlier, Krathwohl *et al* reported their findings on the use of the term ‘attitude’ as follows:

[It was] found to include objectives with a wide range of behaviors. On the one hand, it is used to describe the involvement of the student who is willing to grant that he has a positive feeling about something when he is asked about it. At the other extreme, it is expected that his commitment is such that

¹⁰⁰ A. Lewy, “The Empirical Validity of Major Properties of a Taxonomy of Affective Educational Objectives”, *The Journal of Experimental Education*, (1968), p. 71.

¹⁰¹ R. Colwell, *The Evaluation of Music Teaching and Learning*, (1970), p. 177.

¹⁰² D. Krathwohl, B. Bloom, and B. Masia, (1964), *op. cit.*, p. 37.

¹⁰³ *Ibid*, p. 38.

he goes out of his way to express it and even seeks instances in which he can communicate it to others. Objectives dealing with attitudes frequently require the individual to have a clear conception of his attitude which he can verbalize. ¹⁰⁴

The other terms adopted by Krathwohl et al for inclusion in their hierarchy chart were:

- i) Interest and Appreciation: both of which involve a basic level of a person merely being 'aware' of a phenomena, through to a level whereby the person actively seeks out that phenomena and is captivated by it,
- ii) Value: which covers the same range on the continuum as attitude, but which may be taken as being a more 'umbrella' term - for example, a related 'cluster' of attitudes forms a 'value complex', and
- iii) Adjustment: which is an advanced state of affective development, involving values being interrelated in a more permanent and complex fashion. The individual's whole outlook and behavior may be influenced according to the extent of their 'adjustment' level.

The call, from earlier discussion, is obviously for well-grounded, logical principles on which to base affective domain assessment procedures, hence the decision of the present study to work with the taxonomic model developed by Krathwohl *et al.*. The proposal is to base a music attitude assessment instrument on the Taxonomy and examine the outcome, with a view to meeting the challenges associated with the affective domain and music education.

It is appropriate at this point to acknowledge the fact that traditional procedures for developing attitude scales frequently employ such psychology-based procedures as Thurstone's method of 'equal-appearing intervals' [see in particular the procedural discussion of Kate Hevner's 'Tests for Attitude Toward Music', pp.73-75]. However, as far as the present study is concerned, such methods are being by-passed in favour of

¹⁰⁴ Ibid, p. 25.

examining fully the implications of assessing attitude [in this case, to music] according to a pre-determined, well developed *theory* of affective development (as opposed to a method which revolves around a 'judging-group'-generated value system - see footnote no. 185). Undertaking investigation along these lines is not unknown [see, for example, Lewy (1968), Purcell (1968) and Noble (1982), in Chapter IV]; and it is intended here to add to this body of research whilst at the same time addressing an issue relevant to New Zealand secondary school general music education.

2. NEW DIRECTIONS FOR THE ASSESSMENT OF ELEMENTS IN THE AFFECTIVE DOMAIN

As far as the affective domain and music education is concerned, teachers need not only to be able to 1) compile evidence of the affective growth of students, but also 2) simply obtain details of the 'status quo'¹⁰⁵ of the students; in other words, they must have detailed information as to the nature of the affective relationship that members of a class have with music. The word 'music' here - particularly if one is concerned with the attitudes of adolescent-age students - must of course embrace all genres (for example, from popular to classical styles) and take into account any structured or casual involvement in musical pursuits (including community, school or ethnic activities). Merely 'assuming' the attitudes or levels of musical involvement of a class is not sufficient to ascertain the 'status quo' mentioned above - and in fact may be quite misleading.

Bullock, in his thesis, suggests:

If a junior college instructor assumes that his music appreciation class has an open attitude toward serious music, he is likely to lose touch with the class and fail to interpret music and musical experiences in a manner "relevant" to the students. The frequent result of such failure is student alienation, which

¹⁰⁵ L. Long, "A New Type of Music Interest Scale", *Etude*, (1948), p. 18.

*leads in turn to a strengthening of negative attitudes and contempt rather than appreciation.*¹⁰⁶

Udinsky *et al* identify an 'expectancy' or 'pygmalion' effect, which is how they describe the situation of a researcher who directs his attention to seeing only what he "expects to see".¹⁰⁷ They extend this definition to education, painting the scenario of a teacher who "...is told by the principal that one of her classrooms has all gifted children and another has all slow children; the teacher interprets a noisy atmosphere as 'creativity' in one room and 'disruption' in the other."¹⁰⁸ The knowledge of the cognitive abilities of the class has exerted a great influence on the teacher who works from this information and interprets the behavior of the students accordingly. It is conceivable that the 'gifted' children will receive work that is challenging and stimulating, whereas energies directed at the 'slow' children will be of a more 'controlling' nature.

In short, it is possible that students are being neglected or simply 'written off' musically because they appear to show no aptitude for music. These students may in fact possess an advanced affective relationship with music, but it is one which is not [due, for example, to various environmental reasons] indicated by their cognitive/intellectual level. [It will be recalled that some studies suggest there is little relationship between affective and cognitive growth (p.30)]. Traditional assessment methods, being cognitively based, do not favour affective development, so it is important specifically focus on the latter during the assessment process.

If one supports the view that general music education is concerned with fostering, for example, creativity, experimentation and affective development, then one must welcome

¹⁰⁶ W. Bullock, Construction and Evaluation of a Test of Musico-Aesthetic Attitude, (1971), p. 4.

¹⁰⁷ B. Udinsky, S. Osterlind and S. Lynch, Evaluation Resource Handbook: Gathering, Analyzing, Reporting Data, (1981), p. 93.

¹⁰⁸ Ibid.

assessment procedures that do not stifle such endeavours. What is needed (certainly in New Zealand secondary school general music education) is flexibility to cater for the needs of individuals whilst being aware of the needs of groups - with regard to minimising ethnic and cultural biases. In the New Zealand Department of Education's Syllabus For Schools: Music Education - Early Childhood to Form Seven (1989), the following recommendation is included:

The criteria selected for assessment of students *should be suited to their age, stage of musical development, cultural background, and learning level.*¹⁰⁹

However, the desirability of 'flexibility' in assessment procedures is not to be confused with informality and unstructured subjectivity, as these concepts serve only to undermine the overall efficiency and effectiveness of the educational process.

It was suggested above that "the call is ... for well-grounded, logical principles on which to base affective domain assessment procedures" (p.40), hence an aim of the present study was "to work with the taxonomic model developed by Krathwohl *et al.*" (p.40)

The following discussion considers the points that were raised in Chapter 1 regarding challenges associated with affective domain assessment and music education. Suggestions are made in appropriate instances of possible solutions which may be incorporated into an attitude assessment instrument.

(1) The lack of constructive research:

An examination of the literature has highlighted a need for further constructive research into affective domain assessment; it has been discovered that the body of research into the

¹⁰⁹ Syllabus for Schools: Music Education - Early Childhood to Form Seven, New Zealand Department of Education, (1989), p. 10.

evaluation of cognitive objectives far outweighs that of affective objectives. The current study has been undertaken partly in response to the need to address that imbalance.

(2) The element of subjectivity in affective domain assessment:

By basing an affective domain assessment instrument on the taxonomy of Krathwohl *et al*, one is lending 'educational validity' and accountability to the assessment procedure. The logical principles underlying the taxonomy structure will be identifiable in an assessment instrument developed according to those principles. R. Dave writes:

*A sound evaluation system should be based on instructional objectives. The taxonomies [both of the cognitive and the affective domains] ... provide a logical classification of these objectives, and hence they provide a logical classification scheme for evaluation material based on objectives.*¹¹⁰

(3) The requirement for assessors to adopt flexible standards:

As was indicated in Chapter 1, one of the disadvantages of affective domain assessment is not being able to employ 'right and wrong' policies. Educationalists must resist the urge to dictate 'good' and 'bad', as far as personal taste and experience is concerned; hence the need for developing student-dictated criteria.

One may identify existing attitude (and other affective elements) scales which are now either outdated in themselves, or are based on outmoded concepts [for example Lewy (1968); Noble (1976); and Long (1948)]¹¹¹. There is usually a hang-up on 'good' music ('good' often being synonymous with 'classical'), which in itself is a culturally-biased concept. [Farnsworth, in his "Rating Scales for Musical Interests"¹¹², displayed a

¹¹⁰ Dave, (1969), op. cit., p. 212.

¹¹¹ These will be discussed in Chapter IV.

¹¹² P. Farnsworth, "Rating Scales for Musical Interests", The Journal of Psychology, (1949).

credible receptivity to the notion that an extremely broad pool of musical interests is worthy of evaluation.]

One needs to be sharply aware of such an issue as cultural bias, for unless cultural differences are at least recognised in educational policies, then inflexible assessment procedures will continue to favour the dominant culture. The Encyclopedia of Educational Evaluation maintains that in order for a test to be “culture-fair”, it should:

*reflect comparable, rather than nonexistent, influences from various cultures, so that the test taker is not penalized because of his immersion in any particular culture or lack of opportunity to acquire certain other cultural experiences.*¹¹³

Questions of cultural bias are of course not restricted to differing ‘ethnic cultures’, but also relate to differing age-determined cultures in society. Attention must be drawn here to the relevance to this study of issues concerning assessment processes and the adolescent sub-culture.¹¹⁴

In soliciting the “detailed information as to the nature of the affective relationship that members of a class have with music” (p.41), an assessment instrument must interpret the concept of ‘music’ as widely as is possible [see definition, p.41]; thereby recognizing the values and rights of all cultures and groups.

¹¹³ S. Anderson, S. Ball, R. Murphy and associates, The Encyclopedia of Educational Evaluation, (1975), p. 107.

¹¹⁴ Reference is made here to N. Short, “The Adolescent Sub-Cult: An Issue For Music Education”, The Canterbury Series: Studies in Music Education, (1986), pp 25-33.

(4) The problems associated with the ‘credibility gap’:

The ‘credibility gap’, “refers to the lack of certainty - particularly in the affective domain - ...that a given behavior is evidence of achievement of the desired objective.” [see p.10]

With any affective assessment procedure one needs to be aware of the potential hindrance, and exact location, of the credibility gap factor.

In 1974, Abeles developed the following formula (which was discussed with regard to ‘Affective Measures’ in Foundations of Music Education):

$$BI \approx BH - AS + En$$

where

BI = behavioral intentions, as typically might be obtained from an attitude scale response

BH = behavior, an overt action reflecting the student’s values

AS = affective set, the relatively stable value for an object or event

En = environmental factors, conditions which affect the relationship between the affective set and the observed behavior.¹¹⁵

With regard to the formula Abeles explains: “Behavioral intentions (BI) are approximately equivalent to behaviors (BH), which are due to the individual’s affective set (AS) and environmental factors (En).”¹¹⁶ As far as the present study is concerned, one would be aiming to diminish the impact of any strongly detrimental environmental factors. [See below, under ‘(5) The issue of test anxiety in affective domain assessment’, and ‘(6) The anonymity issue in affective domain assessment’ for suggestions.] In this way, students’

¹¹⁵ H. Abeles, C. Hoffer and R. Klotman, Foundations of Music Education, (1984), p. 247.

¹¹⁶ Ibid, pp. 247-248.

indicated behavioral intentions (for example, their responses to an affective assessment instrument) would approximate as closely as possible to their affective set. Incidentally, the word 'environmental' from the formula is applied somewhat loosely in the context of the present study. For instance, the anonymity issue, whilst not ostensibly fitting the criteria of an environmental factor, is certainly a condition which, along with the issue of test anxiety, affects 'the relationship between the affective set and the observed behavior'. The word 'behavior' is taken by the present study to signify an individual's conduct, or way of acting, that may be examined with respect to an item statement on an attitude inventory. It is not related to the 'behaviorist' movement in the psychological sense [see p.24].

(5) The issue of test anxiety in affective domain assessment:

It was suggested that with the problem of test anxiety, "the solution may begin with seeking to minimise the negative influences at work on the psyches of the students." (p.22)

Solutions here, which may be incorporated into an assessment instrument, would focus on removing the fear of conventional testing. Ways in which this may be achieved would include the following:

- (a) do not insist that a class of students always be assessed at the same time (for example, testing over the period of a week would lessen any negative psychological effect of the instrument),
- (b) avoid imposing a time limit for completing the instrument, and
- (c) allow the respondents to seek comprehension assistance from the teacher at any point.

An attitude assessment instrument taking into account the above factors (factors which allow for a more relaxed and informal approach) would certainly minimise conventional test anxiety.

(6) The anonymity issue in affective domain assessment:

The issue of anonymity, it will be recalled, concerns the question of whether students' performances in an affective domain assessment process "...are influenced by the fact that they are required to identify themselves." (p.11) Variables include:

- (a) the stress of a testing situation (see above, under '(5) The issue of test anxiety...`'), and
- (b) the inappropriateness of 'prior knowledge' of affective assessment objectives (hence the opportunity arising for students to manipulate their responses in an effort to make the 'right' impression. Recall here the issue of the 'credibility gap').

As far as (b) is concerned, the main aim for an attitude assessment instrument would be for the respondents to be unaware, as much as possible, of the underlying reasons for the test. In this respect, the instrument would be operating on two levels: i) a 'basic' level; one that the respondent would understand and respond to, and ii) a 'disguised' level; one that was, in a sense, hidden from the respondent, but which provided (by inference) the information required by the assessor.

An example will help to make the relationship between anonymity and affective domain assessment - with regard to two 'levels' - more clear:

The question "Do you like music at school?", when asked by the music teacher, is more likely to be answered candidly by a student if anonymity is allowed. On the other hand, if the teacher were to ask a number of questions pertaining to various aspects of school music, he would be able to glean, by inference, a measure of the student's overall attitude. Suitable questions could refer to attitudes to types of music covered, and activities undertaken in class.

- (7) **A lack of clearcut definitions with regard to objectives statements, and the confusion arising from the multiplicity of terms employed in the literature:**

All too often affective objectives for a general music curriculum are extremely vague; for example, in the New Zealand Department of Education's music education syllabus, it is stated that students should be 'led' to "contemplate and appreciate music through responsive listening", and to "develop the urge to explore and experiment with music."¹¹⁷ Such objectives are of almost no help to a teacher wishing to develop appropriate teaching and assessment methods, as little or no guidance is given as to how to 'lead' the students in this respect. The taxonomy of Krathwohl *et al*, however, offers practical assistance in these instances, as Lewy suggests:

*The affective taxonomy makes it easier to state affective educational objectives with a high degree of specificity. To illustrate, an objective of the form "to appreciate good music" gives very few hints to the teacher as to what behavior may be expected on the part of the student. Formulating the objectives in a way that reflects its taxonomic level might help to eliminate ambiguities of this sort.*¹¹⁸

Being more specific, R.Dave writes:

Test items are often constructed without conscious effort to base them on predetermined instructional objectives. Emphasis on form and subject matter in constructing a test item, without emphasizing the objective, is likely to result in information-ridden test material. Taxonomic classification alerts the item writer to take into consideration this important base of an item. A

¹¹⁷ Syllabus for Schools: Music Education - Early Childhood to Form Seven, (1989), op. cit., p. 6

¹¹⁸ A. Lewy, "Affective Outcomes of Musical Education", Journal of Research in Music Education, (1971), p. 365.

*conscious and skilful attempt to base an item on a selected objective would help in building content validity in the test material.*¹¹⁹

(8) The expectations of the students themselves, as far as affective measurement is concerned:

The essence of this problem is that educationalists often have to contend with a lack of history of assessment of affective objectives:

Students expect to have their knowledge and their comprehension skills examined, but not their attitudes, feelings, interests, and values [see p.15].

Although one must not think in terms of ‘competition’ with cognitive measures, the profile of affective domain assessment needs to be heightened.

By utilising a ‘structured’ assessment instrument, the teacher would be employing an assessment procedure that would be acceptable, in traditional terms, to the students. The resulting ‘scores’ from such an instrument would furnish concrete evidence of aspects of a student’s affective development.

While this may be seen as contradicting the goal of minimising test anxiety, it is argued that the suggestions given on page 47 remain valid.

(9) The ‘personal’ and private nature of the affective domain:

When undertaking assessment in the affective domain one needs to ensure that the information being solicited is not excessively personal to the respondent, and that the assessment procedure itself is non-threatening. [See above, under ‘(5) The issue of test anxiety in affective domain assessment’ for suggestions.]

¹¹⁹ Dave, (1969), op. cit., p. 213.

Particularly, too, in the case of adolescent education, it should not be necessary to publicise any results of affective domain assessment: one must be aware of the possibility - or even likelihood - of peer group pressure influencing a student's performance. As Charles Hoffer suggests:

*The need for acceptance by one's peers is very strong during adolescence, and accounts for the fact that teenagers conform to fads in dress, grooming, and behavior, even when they don't particularly like or agree with these fads.*¹²⁰

(10) The cognitive versus affective domains, and associated problems of categorisation:

The solution to this problem is to employ a model such as the affective taxonomy of Krathwohl *et al* for defining affective objectives, and for formulating assessment guidelines reflecting those objectives. The extent to which the taxonomy defines the affective domain may be used to advantage in the construction of an attitude assessment instrument.

In short, one is looking to develop an attitude assessment instrument that:

- (a) will furnish details of the 'status quo' of students' affective relationships with music,
- (b) is not drawing primarily from the cognitive/intellectual abilities of students,
- (c) reflects the contemporary call for formative evaluation,
- (d) embraces the framework of Krathwohl's taxonomy of affective objectives,
- (e) is nonthreatening for respondents, with a view to minimising detrimental environmental factors, and
- (f) addresses the issue of cultural/social unfairness.

¹²⁰ C. Hoffer, Teaching Music in the Secondary Schools 2 ed, (1973), p. 128.

CHAPTER III - THE AFFECTIVE COMPONENT OF NEW ZEALAND SECONDARY SCHOOL GENERAL MUSIC EDUCATION

It is necessary to now focus on general music education as it applies to New Zealand secondary schools. The objective will be to identify the affective component of modern-day New Zealand secondary school music education as it claims to be through syllabi - thereby describing the circumstances which frame the current quest to develop an attitude assessment instrument.

By way of a background to this discussion, it will be useful to briefly trace the history of the New Zealand secondary school system, since the late 1930's, in particular with regard to the development of policies and aims.

1. THE SECONDARY SCHOOL IN NEW ZEALAND¹²¹

A starting point for this examination is H.G.R.Mason's article "Some Broad Tendencies", first published in 1944.¹²² Mason pointed out that the state school system was initially "not intended to cater at all levels for the whole population, nor," he stated, "was it intended that schools should concern themselves with the education of every side of the pupil's being."¹²³ Mason drew attention to a report given in 1939 by the Rt. Hon. Peter Fraser (then Minister of Education and Health). An important statement from this report (which was quoted also in James Thorn's biography Peter Fraser, New Zealand's Wartime Prime Minister)¹²⁴ indicated a change in educational policy:

¹²¹ Segments of this Chapter originate from the Author's honours dissertation, An Approach to General Music Education in New Zealand Secondary Schools, (1986).

¹²² In Education Today and Tomorrow, (1944).

¹²³ H. Mason, "Some Broad Tendencies", in Schools in New Zealand Society - A Book of Readings, eds: G. Robinson and B. O'Rourke, (1980), p. 42.

¹²⁴ (1952).

*The government's objective, broadly expressed, is that every person, whatever his level of academic ability, whether he be rich or poor, whether he live in town or country, has a right, as a citizen, to a free education of the kind for which he is best fitted, and to the fullest extent of his powers.*¹²⁵

Fraser was keen to ensure that post-primary education was not restricted to the "well-to-do or the academically able"¹²⁶, but he felt too, that it was not enough to merely remove the selective entry principle. In his view:

*Schools that are to cater for the whole population must offer courses that are as rich and varied as are the needs and abilities of the children who enter them: this means generous equipment, more and better trained teachers, and some system of guidance to help pupils to select the schools and courses that will best cater for their abilities.*¹²⁷

Thus the school system appeared to be taking a wider, more comprehensive role: one which contrasted with the earlier principle that "...assumed that the home, the church, and the job would continue to provide that broader education of the whole being - mental, moral, and physical - which they had been giving in one form or another for centuries."¹²⁸

¹²⁵ Mason, (1980), op. cit., p. 43.

¹²⁶ Ibid.

¹²⁷ Ibid.

¹²⁸ Ibid, p. 42.

In 1942, Mason (quoted above), who was Minister of Education, set up a consultative committee to discuss specific new developments in the New Zealand secondary school system. The ensuing report of this committee, known as the “Thomas Report”, set out as one of its main objectives that:

*...all post-primary pupils, irrespective of their varying abilities and their varying occupational ambitions [should] receive a generous and well-balanced education.*¹²⁹

The report continued:

*Such an education would aim, firstly, at the full development of the adolescent as a person; and, secondly, at preparing him for an active place in our New Zealand society as worker, neighbour, homemaker, and citizen.*¹³⁰

An interesting point to note at this stage is the use of the term “adolescent” in this statement, perhaps indicating a growing recognition that post-primary school pupils form a distinct educational group.

A 1976 Committee on Secondary Education Report, Towards Partnership (See p.58), says of the “Thomas Report” that its “recommendations...have largely determined the present intention, organisation and curriculum of secondary schools.”¹³¹ One of those recommendations -

A school that takes the hard road will re-examine its whole theory and practice, make up its mind about the real needs of its pupils and the means

¹²⁹ The Post-Primary School Curriculum, Report of the Committee appointed by the Minister of Education in November, 1942, “The Thomas Report”, (1959), p. 5.

¹³⁰ Ibid.

¹³¹ Towards Partnership, Report of the Committee on Secondary Education, New Zealand Department of Education, (1976), p. 12.

by which they can best be met, and then act courageously in accordance with its findings ¹³²

- certainly heralded a new era of self-examination by those involved with educational policy-making. One might suggest that this progression reflected a move to reconcile New Zealand education with the needs of New Zealand people; schools hitherto having been modelled largely on their English counterparts.

The Towards Partnership report also made mention of the 1962 Report of the Commission on Education in New Zealand, stating:

*The Currie [committee chairman] Commission asserted that the physical, moral, emotional and intellectual aspects of a student's life are interdependent and all of them are partly the concern of the school.*¹³³

Towards Partnership added, however, that:

*...the commission made it clear that it saw the primary purpose of New Zealand schools as the intellectual development of each student to his full capacity.*¹³⁴

Viewing the policies of the Currie Commission through the eyes of Towards Partnership, there would appear to be a merging of pre-1940s' values with those of the post-Thomas Report era. In fact the Commission had a more contemporary interpretation of the term "intellectual development" than Towards Partnership allowed to show, viz;

*...intellectual development must mean also the cultivation, to the appropriate degree, of aptitudes and attitudes of mind, the ability to think, communicate, judge, and discriminate.*¹³⁵

¹³² The Post-Primary School Curriculum, (1959), op. cit., p. 4.

¹³³ Towards Partnership, (1976), op. cit., p. 12.

¹³⁴ Ibid.

¹³⁵ Report of the Commission on Education in New Zealand, (1962), p. 21.

In 1969 the Curriculum Review Group of the New Zealand Post-Primary Teachers' Association compiled the report Education in Change. The main focus of this report was to identify guidelines for evaluating changing teaching practices. The committee's enquiry included assessment of educational aims and objectives, with a view to allying secondary school education with the growth-process of adolescence. Emphasis was placed on promoting in the students, "the urge to enquire; concern for others;" and "the desire for self-respect."¹³⁶ These qualities were manifested in educational aims that were summarised by the committee as follows:

- (1) To help young people to acquire values which will aid their individual growth and social consciousness.
- (2) To develop their capacity to enquire in representative fields of study.
- (3) To develop their ability to relate to the parts of their educational experience so that they can see this experience as relevant to their lives.¹³⁷

The third of these aims indicated that educationalists were beginning to think of schools in terms of being accountable to the adolescents themselves, with regard to their non-school lives. It was perhaps this type of concept that gave rise to such programmes as the 'Transition To Work' and 'Work Experience' schemes that have so coloured subsequent educational developments.

The role of the secondary school in New Zealand society was broadening still.

¹³⁶ Education in Change, Report of the Curriculum Review Group, New Zealand Post Primary Teachers' Association, (1969), p. 1.

¹³⁷ Ibid, p. 71.

In 1971, the Department of Education conducted a course for secondary and tertiary educators that centred around three major issues, as follows:

- (1) The knowledge, skills, attitudes and values we should be developing in students;
- (2) The functions of secondary schools; and
- (3) The nature of a general education.¹³⁸

One policy agreed upon regarding the second issue was that “Schools should be organised to provide for the intellectual, social, emotional and physical growth of each student.”¹³⁹

This was further expanded by one group that decided that:

- (1) Schools are for preparation for:
 - living in society;
 - further learning;
 - assisting students to adjust to change;
 - self-realisation.
- (2) Schools are for living in and students are entitled to:
 - security;
 - recognition as individuals and as members of groups.
- (3) Schools are for intellectual stimulation.¹⁴⁰

It is interesting that, in the above view, “intellectual stimulation” was not the sole primary objective: compare with the emphasis of the Currie Commission nearly ten years before [refer p.55]. Also worthy of note is the absence of specific reference to job preparation.

¹³⁸ The Secondary School Curriculum: Some Issues and Prospects, New Zealand Department of Education, (1971), p. 1.

¹³⁹ Ibid, p. 10.

¹⁴⁰ Ibid, p. 12.

During the early 1970s a review of New Zealand secondary education was set in motion, with participation sought from every secondary school in the country. The initial purpose was to discuss “the nature and purpose of secondary education”, which then grew to incorporate an examination of “aims and activities”.¹⁴¹ In 1975 the Minister of Education established a national committee to consider recommendations from the reports of the various area meetings. Results of the committee’s findings were published in the report Towards Partnership in 1976. In this report it was stated that “the education provided in our secondary schools should be based on and should exemplify the values of mutual concern, individuality, co-operation, autonomy, adaptability [and] quality.”¹⁴² In examining the issue, ‘Aims of Secondary Schools’, the report outlined the problem of “no generally agreed and explicit statement of aims for New Zealand secondary education.”¹⁴³ A result of this problem, it was felt, was that external examinations and qualification-seeking were receiving too great an emphasis. The report, therefore, recommended that “each secondary school formulate and periodically review its own aims in consultation with parents and students.”¹⁴⁴ This recommendation echoed one made in the 1971 Department of Education publication [refer p.57] Here it was stated:

*Aims and needs of society change from time to time and these are seldom clearly stated. The school has a responsibility to specify its objectives and the values on which they are based. Schools have a continuing responsibility to evaluate their success in achieving their objectives.*¹⁴⁵

By formulating its own aims and policies, the secondary school may decide for itself how it can best serve the needs of the immediate district.

¹⁴¹ Towards Partnership, (1976), op. cit., p. 13.

¹⁴² Ibid, p. 5.

¹⁴³ Ibid, p. 7.

¹⁴⁴ Ibid.

¹⁴⁵ The Secondary School Curriculum: Some Issues and Prospects, (1971), op. cit., p. 10.

Having discussed the history of the New Zealand secondary school system from the late 1930s, it is now time to examine the corresponding development of music education.

2. THE DEVELOPMENT OF CONTEMPORARY NEW ZEALAND MUSIC EDUCATION

In the early part of this century, New Zealand music education was very firmly steeped in the English tradition. Douglas Tayler, an 'imported' Englishman, wrote A Scheme of School Music Related to Human Life in 1927, whilst he was 'Supervisor of Musical Education to the New Zealand Government'. The approach he advocated in this publication, under "Syllabus of Instruction", embraced "...the training of the voice and of the ear; the correlation of music with rhythmical and interpretative movement; the teaching of musical theory, and sight-singing, including use of the modulator; encouragement of musical invention; songs to be studied; and suggestions for musical appreciation lessons, and the use of the gramophone."¹⁴⁶ Tayler had high aspirations for the teaching of music in schools, due to his beliefs that music "...springs spontaneously from the overflowing of happiness, and so recreates happiness when heard"; that "beautiful music stimulates the emotions and the imagination in beautiful ways, leading to the creation of beauty in life"; and also that music "...is a pleasant, wholesome, social recreation."¹⁴⁷ The emphasis was, even so, primarily on skill acquisition, which then purportedly led to the enjoyment of music:

*All theory must be regarded as the necessary equipment...and by teaching it through construction of music as well as through analysis, a much greater interest is maintained.*¹⁴⁸

¹⁴⁶ E. Douglas Tayler, A Scheme of School Music Related to Human Life, (1927), p. 15.

¹⁴⁷ Ibid, p. 7.

¹⁴⁸ Ibid.

In 1941, Vernon Griffiths wrote that “‘school music’... should be regarded as a basic activity at least equal in importance to the traditional school subjects.”¹⁴⁹ Griffiths’ objective for music education in schools incorporated the hope that ‘lasting enthusiasm’ for music could be generated in the pupils, resulting in “active participation...as performer, listener, or creator.”¹⁵⁰ Griffiths’ main aim was to develop a totally musically active school; the emphasis being - even more specifically than with Tayler - on performance and vocal skill development.

The somewhat ‘cut-and-dried’ approach of music education in the 1920s and 1930s was commensurate with the educational policy of the time, that being, the prime importance of academic learning, and the advancement of intellectual skills. H.G.R. Mason made the observation in 1944, that there was a new “growing recognition of individual differences between children”. One realises that, prior to this, students were expected to conform unquestioningly to the requirements of a subject. One philosophy behind this concept was suggested by Fisher, in the late 1940s. He wrote:

Until recently it was a cardinal principle of education that drudgery in schools fortifies the character and enables the child to adapt itself later to the inevitable drudgeries of life. At the present moment our schools embody both this and the more recent conception of a curriculum based on the interests of the children; of school work as a happy and absorbing activity, developing the child’s personality and personal aptitudes, rather than moulding the child to the curriculum. ¹⁵¹

¹⁴⁹ T.V. Griffiths, An Experiment in School Music Making, (1941), p. 1.

¹⁵⁰ Ibid.

¹⁵¹ N. Fisher, “Music and the Curriculum of the Future”, in Musical Education, ed: H. Watkins Shaw, (c. 1945), p. 71.

An interesting aside here, is to refer to Christopher Small who, in 1977, still could not share the optimism of Fisher on this point:

*No child, at least after his first few years at school, ... actually expects school to be an enjoyable experience; if it proves so, he regards it, rightly, as a lucky bonus. He is made aware that what he is being subjected to is in his best interest, for the sake of some vaguely-defined future advantage which for most remains perpetually out of reach.*¹⁵²

Returning now to the 'Thomas Report' (1942), it will be remembered that the new aims recommended for education included provision for "a generous and well-balanced education", and also "full development of the adolescent as a person" [refer p.54]. These aims could be seen perhaps as the secondary school environment gradually becoming receptive to a more flexible approach to music education; more specifically, an approach that did not have academic and skill development as its ultimate objective. And indeed, the Report's section entitled "Music, Arts and Crafts" stated: "We feel that a definite place should be made in the timetable for aesthetic activities."¹⁵³ As to their standpoint on the subject of aesthetic education in secondary schools, the Report Committee quoted the 1935 "Spens Report", stating that it recommends "Music and the Arts 'for their value in awakening and developing that aesthetic sensibility which is one of the most valuable of human gifts, and which, although its possibilities vary greatly from one child to another, is wholly denied to none.'"¹⁵⁴

It would seem then that New Zealand music education, as directed by the 'Thomas Report' thus far, was moving in positive directions and adapting well to changing educational policies. Unfortunately, further examination reveals otherwise. The "well considered

¹⁵² C. Small, Music-Society-Education, (1977), p. 184.

¹⁵³ The Post-Primary School Curriculum, (1959), op. cit., p. 40.

¹⁵⁴ Ibid, p. 41.

programme of activity”¹⁵⁵ favoured by the Report included the teaching of sight reading, voice production, and “the elementary grammar of music”; instrumental programmes à la Vernon Griffiths; and the suggestion that “short talks on music should frequently be given” as the “Music as Appreciation” component. Aesthetic education as a policy received no further treatment; instead, traditional objectives were adapted, for example: “The choir, which may well contain all the pupils of the school, should get continual practice in the morning assembly, and should play a large part in enhancing the corporate life of the school.”¹⁵⁶

The fact remains, nevertheless, that throughout the history of New Zealand secondary school education, it has been accepted that music is an integral part of a good general education [“core music” becoming compulsory in 1944].

3. THE NEW ZEALAND MUSIC EDUCATION REVIEW

The form of general music education in New Zealand remained relatively static for nearly forty years - until the advent of the “Music Education Review” in the 1980’s.

In an address delivered to the 1985 conference of the New Zealand Society for Music Education, the then Director- General of Education, Mr William L. Renwick, outlined the progress of the much overdue “Music Education Review”. Renwick reported that he was originally authorised “to convene a meeting of music teachers to plan a comprehensive review of music education.”¹⁵⁷ In his words:

We met in 1984 and, in August 1984, the present Minister accepted our recommendations and set up the National Music Review Committee. It is a widely representative and very experienced committee. It includes practising

¹⁵⁵ Ibid, p. 42.

¹⁵⁶ Ibid.

*teachers from pre-school to university, music advisers and teachers college lecturers, state and private teachers organisations, the School Certificate Examination Board and the Universities Entrance Board, the Institute of Registered Music Teachers, Radio New Zealand, leading Maori, Samoan and Cook Island teachers, and officers of the department.*¹⁵⁸

Renwick drew attention to the fact that New Zealand music per se was now playing an important role in the establishment of a national, cultural identity. He added:

*...the creative expression of New Zealanders is at last coming to be valued. ...More and more New Zealanders are at last taking it for granted that creative expression is a normal feature of our cultural life. Even more important, they are coming to understand that a cultural life that is not rooted in the creative expression of its own people is derivative and inauthentic.*¹⁵⁹

The 'new look' music syllabus aimed to reflect the nature of contemporary New Zealand society by highlighting the multicultural perspective. In this way, the Review group was endeavouring to break away from the constraints of the colonially-imported English approach to music education:

...the Music Review Committee is giving a broad meaning to the phrase musical culture. It is following the general contemporary usage which takes culture to refer to the shared meanings that the members of a community have in common. From this it follows that the task of music education is to introduce young New Zealanders into the best manifestations of the various forms of expressions of our musical expression. New Zealand music, as the Music Review Group is using the phrase, thus refers to music composed or

¹⁵⁷ W. Renwick, "Opening Address" in New Directions in Music Education, Year Book 1985, New Zealand Society for Music Education (Inc), (1986), p. 9.

¹⁵⁸ Ibid.

¹⁵⁹ Ibid, p. 10.

*created by New Zealanders, in any style, and includes electronic music, rock, jazz, folk music, Maori, Samoan, Tongan, Cook Island, Tokelauan, Niuean and the music of any other cultural group living in this country whose music is part of their expression as a people.*¹⁶⁰

By way of summarising the work of the Music Review Committee, Renwick listed the following “main conceptual issues”:

*music as creative activity; musical creation as a vital aspect of our developing cultural identity as New Zealanders; the diverse forms of musical expression that are important to New Zealanders and what must be done to reflect that diversity in music education; and music in relation to other art forms.*¹⁶¹

In 1987 the Draft Syllabus for Music Education: Early Childhood to Form Seven was completed and distributed, along with a questionnaire designed to assess the impact of the document:

*This draft syllabus has resulted from discussions during the last two and a half years. It is now released to enable teachers and other interested people in the community to consider its appropriateness as a basis for music education in schools.*¹⁶²

¹⁶⁰ Ibid, p. 11.

¹⁶¹ Ibid, p. 14.

¹⁶² Draft Syllabus for Music Education: Early Childhood to Form Seven, New Zealand Department of Education, (1987), pp. 1-2.

4. THE SYLLABUS FOR SCHOOLS. MUSIC EDUCATION - EARLY CHILDHOOD TO FORM SEVEN

The final form of the new music education syllabus (published and approved in 1989 by the Minister of Education) contains the following extract in the Foreword:

This syllabus provides the basis for music education from early childhood to form 7. A national music review was established by the authority of the Minister of Education in August, 1984, to review music education in all its aspects. A National Syllabus Committee and thirty-two regional committees were also set up. Position papers on fifteen of the most significant topics were compiled by these local committees and were considered by the Syllabus Committee who were very appreciative of the contributions received from teachers and others throughout the country. During the period of review there has been wide consultation, through national and local conferences, with the music advisory service, teachers' college music staff, inspectors of schools, practising teachers of music in schools, members of the Maori and Pacific Islands communities, representatives of NZEI, PPTA, and the examining boards, the Institute of Registered Music Teachers, and other interested groups.¹⁶³

Under the heading, 'Music in Life and Education', the 'rationale' for general music education is summarised:

Among the most significant reasons for including music in the education of all students are that it will enable students to:

- *grow aesthetically through musical experiences;*
- *realise their musical potential;*

¹⁶³ Syllabus for Schools: Music Education - Early Childhood to Form Seven, New Zealand Department of Education, (1989), p. 3.

- *develop important cognitive processes, such as imagining and lateral thinking;*
- *consider the music industry as a viable profession;*
- *gain pleasure and, in many cases, develop a lifelong interest;*
- *become aware of the distinctive functions of music in our society;*
- *affirm and expand knowledge of their own musical and artistic heritage;*
- *gain a greater understanding of cultural similarities and differences through music;*
- *improve personal identity and positive relationships with others through individual and group participation in music;*
- *develop sensitivity to the quality of sound in the environment.*¹⁶⁴

The inclusion of an affective component in these statements is apparent [i.e. they hint at the desirability of improving attitudes, awareness, interests etc], indicating the Review Committee's consideration of the issue. At this point of the syllabus, however, no particular emphasis is placed on affective domain outcomes. Apart from the fifth statement ("gain pleasure and, in many cases, develop a lifelong interest"), the references are general in nature. Interestingly, the actual 'Aim' of music education in New Zealand, as recommended by the music syllabus, is as follows:

*The aim of music education is to involve people in the active, creative processes of making and listening to music, in ways that promote individual aesthetic growth and fulfilment.*¹⁶⁵

¹⁶⁴ Ibid, p. 5.

¹⁶⁵ Ibid, p. 6.

This aim certainly appears to be upholding affective domain goals; particularly the reference to 'promoting aesthetic fulfilment'.

The syllabus continues by outlining the objectives of music education:

Objectives

Music everywhere depends on people who can:

- create music;
- re-create music;
- appreciate music.

The objectives of this syllabus derive from these basic aspects of music, and seek to lead students to:

create music

- develop the urge to explore and experiment with music;
- foster the ability to improvise, arrange, and compose music;
- make imaginative responses in sound;

re-create music

- enjoy performing music;
- encourage music-making of all kinds, including the re-creation of their own and other students' creative efforts;
- share musical performances;

appreciate music

- develop the ability to identify, describe, classify, compare, and analyse, in order to understand and value music;
- contemplate and appreciate music through responsive listening;
- acknowledge and respect both quality in music and

- performers who excel.¹⁶⁶

Many key words from these objectives statements - such as 'urge', 'imaginative', 'enjoy', 'appreciate', 'value', and 'respect' - contain an affective domain emphasis. In fact, a good balance of cognitive and affective elements is now emerging in the discussion.

However, a question needing to be asked at this point pertains to problems of 'instructional practices' and 'evaluation'.¹⁶⁷ The objectives of the new music education syllabus are certainly meritorious, yet how are they to be followed through in the classroom settings, and how are they to be monitored or assessed? As advised by Krathwohl *et al*, "Until the objective is clearly specified, it can serve only to indicate broad general directions."¹⁶⁸ The syllabus offers a section headed "Monitoring Students' Progress", in which it is stated: "Monitoring is concerned with ascertaining students' attitudes and interests and assessing their achievement. It also includes recording the information obtained and evaluating whether progress is occurring."¹⁶⁹

Affective domain elements are now receiving higher consideration than their cognitive domain counterparts - a fact which is illustrated even further as the discussion progresses:

Monitoring is the measurement of students' achievement, progress, attitudes, and interest in relation to the objectives of a given programme.

Monitoring involves the assessing and recording of:

- *the extent to which the student has a positive, open attitude to a wide range of music;*
- *how involved the student is with music;*

¹⁶⁶ Ibid.

¹⁶⁷ Refer D. Krathwohl, B. Bloom and B. Masia, Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook II: Affective Domain, (1964), p. 64.

¹⁶⁸ Ibid, p. 64.

¹⁶⁹ Syllabus for Schools: Music Education - early Childhood to Form Seven, (1989) op. cit., p. 10.

- *the breadth and depth of each student's interests and how such interests may have changed;*
- *present attainment level (and thereby reveals progress over a specific, generally short, period of time).*¹⁷⁰

The syllabus cautions that the monitoring of student progress must be based on relevant criteria - criteria which "should be suited to students' cultural background, age, stage of musical development, and learning level".¹⁷¹ The assessment procedure advocated by the syllabus is the 'individual profiles' system, as "it is considered to be the most appropriate record of music assessments."¹⁷²

According to the syllabus, a profile "consists of a written statement of progress and achievement in each element of the programme - aural, creative, performance, study, and appreciation - as well as an indication of the student's attitudes and interests."¹⁷³

However, the above statement contains a distinct anomaly: an imbalance has appeared with respect to the discussed objectives for music education, and the assessment details expected for 'student monitoring'. Specifically, three of the four music education outcomes to be 'assessed and recorded'¹⁷⁴ are predominantly affective in nature; whereas in the 'development of a profile' directions, "an indication of the students' attitudes and interests" is apparently all that is required. Furthermore, no suggestions are offered as to the form that this "indication" should take.

Hence the affective component appears to fare badly at the assessment stage.

¹⁷⁰ Ibid.

¹⁷¹ Ibid.

¹⁷² Ibid, p. 11.

¹⁷³ Ibid.

¹⁷⁴ These were listed above.

The purpose of the above discussion was “to identify the affective component of modern-day New Zealand secondary school music education, thereby describing the circumstances which frame the current quest to develop a music attitude assessment instrument.” It has been shown that affective domain objectives are now accorded official recognition in the new music education syllabus for schools; and yet the procedures for their assessment are inadequately addressed.

CHAPTER IV - REVIEW OF THE RESEARCH

The following 'Review of the Research' aims to present a survey of past studies and research associated with the topic of affective assessment in music.

Several areas need to be isolated for examination - particularly with regard to the work of Krathwohl *et al.*¹⁷⁵ This last will be considered separately, and its implications for the present study discussed. But first, a more general approach to this review needs to be adopted, so as to provide a background to the more specific task of basing an affective assessment instrument on the Taxonomy of Krathwohl *et al.*

1. REVIEWS OF RESEARCH INTO ASSESSMENT IN MUSIC

Three significant reviews of research into assessment in music education, which make reference to affective domain elements, appeared in the 1970s. The first of these, "Research in Evaluation in Music Education" by William Whybrew, makes reference to 'attitudes toward music'¹⁷⁶ studies in the wider context of general testing in music. Of more direct interest to the present study are the reports of William Bullock ("A Review of Measures of Musico-Aesthetic Attitude")¹⁷⁷ and Joel Wapnick ("A Review of Research on Attitude and Preference").¹⁷⁸ Bullock structures his review by categorizing the various types of attitude evaluation instruments under the following headings: 'Measures of Perceptive Ability', 'Verbal Measures of Attitudinal Disposition', and 'Tonal Measures of Attitudinal Response' - all of these containing subcategories as well.

¹⁷⁵ D. Krathwohl, B. Bloom and B. Masia, Taxonomy of Educational Objectives. The Classification of Educational Goals - Handbook II: Affective Domain, (1964).

¹⁷⁶ W. Whybrew, "Research in Evaluation in Music Education", Bulletin of the Council for Research in Music Education, (1973), p. 12.

¹⁷⁷ W. Bullock, "A Review of Measures of Musico-Aesthetic Attitude", Journal of Research in Music Education, (1973).

¹⁷⁸ J. Wapnick, "A Review of Research on Attitude and Preference", Bulletin of the Council for Research in Music Education, (1976).

2. AFFECTIVE DOMAIN AND ASSESSMENT IN MUSIC

Assessment in the affective domain is by no means a new concept in music education; hence at this stage of the 'review of the research' it will be of value to present an overview of past explorations into this area.

The following section is concerned with music assessment research containing links with affective domain objectives. For the purposes of this study, however, this section is divided into two parts, the first of which will examine studies and research which are of distinct interest to the present topic. The particular point of reference for these studies is that they are concerned with verbal tests, as opposed to tonal tests. As defined by Bullock in "A Review of Measures of Musico-Aesthetic Attitude":

*Verbal tests do not employ musical stimuli; tonal tests do. Verbal tests demand either retrospective or prophetic valuation because they do not use actual music. ... verbal tests generally measure attitudinal disposition toward music as a concept rather than music as a stimulus. These tests evaluate opinions and interests rather than preferences...*¹⁷⁹

In short, affective tonal tests generally assess affective response only - refer to p.17 for the 'episodic immediacy' discussion - and do not necessarily elicit a subject's longer-term attitude.

The second part will highlight literature and research which, although appearing to be related, contain aspects which preclude their direct applicability.

These last must be identified so as to further delimit the topic in hand.

¹⁷⁹ W. Bullock, (1973), op. cit., p. 336.

(1) Studies and research directly related to ‘The Affective Domain and Assessment in Music’ (as introduced on p.72):

Valuable pioneering work concerning psychological testing and music was undertaken in the 1930s by Kate Hevner, the beginnings of which were reported in two articles, “Tests for Esthetic Appreciation in the Field of Music”¹⁸⁰ (1930), and “A Study of Tests for Appreciation of Music”¹⁸¹ (1931) [see pp.82-83 for discussion]. Further to these articles, Hevner published her treatise “Appreciation of Music and Tests for the Appreciation of Music”¹⁸² (1934). As well as expounding the theory and rationale behind her own ‘Music Discrimination Test’ (which by this time, following the award of a Carnegie Foundation grant, had been “...extended and improved and permanently recorded for phonograph reproduction.”¹⁸³, Hevner also detailed the development of her ‘Tests for Attitude Toward Music’. Her concern in this instance was to address the issue of assessing individuals’ attitudes towards the ‘value’ of music - in a way that was, to some degree, objective. To this end, she turned to the method of constructing psychological attitude tests as devised by L.L. Thurstone, according to the procedure outlined in Seashore and Hevner’s “A Time Saving Device for the Construction of Attitude Scales.”¹⁸⁴ This procedure involved firstly compiling a large pool of item statements expressing opinions about music, and submitting them to a judging group.¹⁸⁵ Hevner reported that 200 statements were initially collected

¹⁸⁰ K. Hevner, “Tests for Esthetic Appreciation in the Field of Music”, Journal of Applied Psychology, (1930).

¹⁸¹ K. Hevner, “A Study of Tests for Appreciation of Music”, Journal of Applied Psychology, (1931).

¹⁸² K. Hevner, “Appreciation of Music and Tests for the Appreciation of Music”, Studies in Appreciation of Art: University of Oregon Publications, (1934).

¹⁸³ Ibid, p. 95.

¹⁸⁴ R. Seashore and K. Hevner, “A Time-Saving Device for the Construction of Attitude Scales”, Journal of Social Psychology, (1933).

¹⁸⁵ Thurstone’s Method of attitude scale construction: “Statements in interviews from individuals relating to the topic under consideration are collected. ... The statements are short and open to only one interpretation. The statements are classified by [a large number of] judges into favourable, neutral or hostile. Each group is further subdivided. The most relevant, least ambiguous set of statements which covers the whole range of values is chosen. The student selects statements which agree with his feelings.” A Dictionary of Education, ed: P. Hills, (1982), p. 273.

(after consulting “various individuals and writings on music”¹⁸⁶), covering opinions “ranging from deep appreciation to indifference or disparagement”¹⁸⁷. One hundred of these statements were subsequently rejected at this point on the basis of their being too lengthy, uninteresting, or ambiguous. The remaining statements were submitted to a judging group of one hundred subjects. The judging group was directed to rate each item statement according to a nine-point scale, ranging from ‘A’ (the highest appreciation of the value of music), through ‘F’ (neutral/noncommittal), to ‘K’ (the strongest depreciation of music). One instruction was to “base your judgements on the best music you know”¹⁸⁸ - which was interestingly vague, but perhaps understandable considering the fact that the experiment predated any controversy surrounding the relative value of ‘popular’ music. Following this step, fifty item statements were retained, according to these criteria:

1. There must be statements in all parts of the scale, and spaced as evenly as possible from 1 to 11.
2. The statements must be clear and unambiguous.
3. The statements about which the raters differed widely in their opinions must be eliminated.¹⁸⁹

The final form of the tests was two arbitrarily divided scales of twenty-five item statements apiece, each statement being assigned a value (calculated from the ‘judging group’ procedure detailed above) of between 1 and 11. A respondent’s score on the scales is calculated by firstly summing the values of the item statements that have been agreed with, and then obtaining the average. “For example, if a subject should check four statements with scale values of 7.8, 6.4, 7.0, and 7.2 respectively, his ‘score’ would be the average of these four values, or 7.1.”¹⁹⁰

Reliability studies were undertaken by submitting the scales to “230 college students in various classes at the University of Minnesota”.¹⁹¹ The reliabilities were calculated by the

¹⁸⁶ K. Hevner, (1934), op. cit., p. 139.

¹⁸⁷ Ibid, pp. 138-139.

¹⁸⁸ Ibid, p. 139.

¹⁸⁹ Ibid, p. 140.

¹⁹⁰ Ibid, p. 141.

¹⁹¹ Ibid.

odd-even correlation, corrected by the Spearman-Brown formula, with the result of the two scales combined being .90. The reliabilities of the halves calculated separately were .79 and .81 respectively.

Hevner felt these results were high enough to indicate the value of the scales for group assessments (in the case of the two 'half' forms), and for the purpose of assessing individuals' attitudes (when the scales were used in conjunction). No validity studies were undertaken, as Hevner felt that "...there is no reason to suspect that ... validity would be anything but satisfactory, since all of the scales previously constructed by this method have given satisfaction in this respect." ¹⁹²

The Hevner 'Tests For Attitude Toward Music' appear in Appendix B.

Of interest to the present study is the 1948 article by Leland Long, entitled simply, "A New Type of Music Interest Scale."¹⁹³ In this, Long described the development of a 'Music Interest Inventory', designed to "...survey [musical] interests and give a picture of the 'status quo' "¹⁹⁴ of his band and orchestra students. Long wished to examine objectively the extent to which the level of interest that a student possessed contributed to their overall success as a music student. He explained:

*Many students who were superior mentally and rated high in pitch, rhythm, and tonal memory were just average members of musical organizations; whereas, a number of students who were merely average in comparative test scores, were doing quite outstanding work. The thought occurred that the reason for this difference, and in general for many differences in attitude in rehearsal and toward home practice, was the degree of interest each student possessed or had developed in his instrumental work.*¹⁹⁵

¹⁹² Ibid, p. 141.

¹⁹³ L. Long, "A New Type of Music Interest Scale", *Etude*, (1948).

¹⁹⁴ Ibid, p. 18.

¹⁹⁵ Ibid.

The resulting inventory “consisted of one hundred participle phrases denoting various types of musical activity”¹⁹⁶, each falling into one of five categories - those being Classical Listening, Classical playing, Swing Listening, Swing Playing, and Related Interests. [Long commented on the inclusion of the ‘Swing’ categories in his instrument: “...the time worn battle between jazz and the classics ... nearly led to the abandonment of the project. Most high school students are unconscionably dedicated to swing, and there was little point comparing swing and classical interests except for the light which it would throw on the latter.” ¹⁹⁷] Students were asked to consider each activity in turn, and then indicate on an answer sheet a ‘like’, ‘dislike’, or ‘indifferent’ response.

Long listed the outcomes of employing such an instrument as the one he devised as being the following:

- (1) it provides an objective basis to prove or disprove previous conclusions based upon observation alone;
- (2) it uncovers new data regarding students’ interests and personalities;
- (3) it affords a broader base and more detailed information than the usual form of questionnaire;
- (4) it permits a mathematical interpretation, revealing comparative strength and breadth of interest by means of a numerical score;
- (5) it affords an opportunity for exercise of critical judgement; and
- (6) it is an instrument of indoctrination in suggesting a number of worth-while musical activities to the student.¹⁹⁸

Long did not elaborate on the above outcomes, nor did he report undertaking any reliability or validity studies for his Instrument. Nevertheless the important factor of his work is that

¹⁹⁶ Ibid.

¹⁹⁷ Ibid.

¹⁹⁸ Ibid, p. 53.

he, like Hevner, was concerned about formalising the assessment of elements emanating from the affective domain.

Farnsworth, in his article "Rating Scales for Musical Interests"¹⁹⁹ (1949), outlined his work developing an instrument to evaluate musical interests and taste. In a similar vein to Long, Farnsworth approached his assignment from as broad a base as possible, as he felt that past interests tests were inclined to "...tap attitudes toward 'serious' or 'art' music and do little or nothing with the various sorts of 'popular' music."²⁰⁰ Farnsworth's instrument took the form of a 'spatial scale'; as Wapnick stated, "This scale is essentially the same as the equal interval scale, with the exception that subjects can respond anywhere along the line, not just at the reference points."²⁰¹ One end of the scale represented the most negative one may feel about a particular type of music, the other end, the most positive. Farnsworth used a judging group [See footnote, p.73] in establishing the points along the continuum, which were used by the respondents as a guide to where they might place a 'check' on the line, thereby indicating where their attitude lay. Five forms of the scale were produced, one each for 'general' music, 'serious' music, 'popular' music, 'hit parade', and 'waltz'. Farnsworth made small validity examinations, enough for him to conclude that "...the tests would seem to have value, at least in group assessments."²⁰²

Moving on chronologically, the next study into the assessment of musical attitudes, with direct relevance to the present research, was reported in 1971 by Edwards and Edwards.²⁰³ In an article entitled "A Scale to Measure Attitudes Toward Music", Edwards and Edwards wrote of the need to develop appropriate measures for the evaluation of student attitudes:

¹⁹⁹ P. Farnsworth, "Rating Scales for Musical Interests", The Journal of Psychology, (1949).

²⁰⁰ Ibid, p. 245.

²⁰¹ J. Wapnick, (1976), op. cit., p. 4.

²⁰² P. Farnsworth, (1949), op. cit., p. 252.

²⁰³ J. Edwards and M. Edwards, "A Scale to Measure Attitudes Toward Music", Journal of Research in Music Education, (1971).

*If music interpretation and appreciation are to be course goals, they should be measured by a reliable instrument. Although attitudes are usually inferred from behavior, observation and measurement of behavior are difficult. Because attitude can be assessed more easily and more precisely with a reliable attitude scale, this study proposed to develop such a scale for use among college students.*²⁰⁴

The scale they constructed consisted of item statements ranked by a judging group according to a 'most favourable' to 'least favourable' format. The statements were each attached to six Likert-type response categories²⁰⁵, ranging from 'strongly agree' to 'strongly disagree'. The scale was administered, and the results were analysed "...according to Goodenough's method of scalogram analysis, resulting in a coefficient of reproducibility"²⁰⁶ [not given in the article]. Edwards and Edwards reported that they subsequently adjusted the scale so as to increase its reliability ("by eliminating those statements with a large error of reproducibility").²⁰⁷ They then proceeded to prepare two forms of the scale from the statements arising from this initial research. "Statements similar in position on the continuum were paired so that the two forms could be made as nearly equivalent as possible."²⁰⁸

After administering the two scales to 100 subjects, coefficients of reproducibility of .788 and .762 were gained for scale 1 and scale 2 respectively. Citing these results, the authors decided that the "statements scale adequately."²⁰⁹ The internal consistency indexes obtained

²⁰⁴ Ibid, p. 229.

²⁰⁵ A Likert-Type item: "An item in a test of interests or attitudes in which a person is invited to indicate the extent to which he agrees or disagrees with each of a number of given statements, usually on a five-point scale, in which one extreme may represent 'strongly agree', the other extreme 'strongly disagree', and the central point a neutral stance." D. Rowntree, A Dictionary of Education, (1981), p. 159.

²⁰⁶ J. Edwards and M. Edwards, (1971), op. cit., pp. 231-232.

²⁰⁷ Ibid, p. 232.

²⁰⁸ Ibid.

²⁰⁹ Ibid, p. 233.

were .809 and .925 respectively. Edwards and Edwards concluded that “the internal consistency indexes are of sufficient magnitude to indicate reliability when the scales are used as groups.”²¹⁰

The authors admitted the limitations of their attitude scale, stating, by way of conclusion, that “it can indicate the range of attitudes within a given class accurately, but cannot be used to quantify the attitudes of an individual”.²¹¹ They suggested that further investigation was “urgently needed”²¹² so as to establish the construct validity of their scale. It is surprising to note that, while Edwards and Edwards were keen to stress the benefits of their scale for group assessments, they did not address the issue of individual testing. Yet they were concerned about elevating the ‘status’ of attitude measurement to the level of ‘achievement’ measurement.²¹³ The actual scales were not presented in the article.

In 1973, Wallace Nolin conducted research with a view to determining “...what influence the less frequent meetings of classes for music instruction had on the attitudinal growth patterns of the students toward their school music experiences.”²¹⁴ In order to test for attitude toward music, Nolin used a previously developed ‘music attitude inventory’ [Oliver Broquist: “A Survey of the Attitudes of 2,594 Wisconsin Elementary School Pupils Toward Their Learning Experiences in Music”, doctoral dissertation, University of Wisconsin, 1961] which he had modified in 1967. The inventory comprised “...incomplete statements about the various component factors of the music program”,²¹⁵ to which the students were to respond with one of the following four answers: 1) I like it, 2) It's O.K., 3) I don't like it, and 4) I don't remember doing it. [Nolin did not offer further details regarding the items of the inventory.] The response sheets were ‘scored’ according to the

²¹⁰ Ibid, p. 232.

²¹¹ Ibid, p. 233.

²¹² Ibid.

²¹³ Ibid, p. 228.

²¹⁴ W. Nolin, “Attitudinal Growth Patterns Toward Elementary School Music Experiences”, *Journal of Research in Music Education*, (1973), p. 124.

²¹⁵ Ibid, p. 125.

values assigned to each particular answer. The validity of the inventory was "...determined by the approval of ... three music teachers"²¹⁶ (perhaps somewhat insufficient), whilst the reliability was calculated by the test-retest method (achieving coefficients ranging from .723 to .812, which Nolin considered to be satisfactory).

Nolin found that the results of his study replicated "similar declines in attitude expressed by students as they grow older, which have been found in previous studies of this type."²¹⁷ He reported:

*Not once did any grade or sex in this study express higher attitudes than corresponding grades and sex groups in the other two similar studies. [Broquist, 1961 and Nolin, 1967] It cannot be said conclusively that less frequent meetings for music instruction alone produced lower expressed attitudes of students. Many other variables could contribute to those lower attitude scores. But likewise it cannot be said, as was conjectured before the study was started, that less frequent class meetings produced higher attitudes.*²¹⁸

Despite the fact that the inventory was only a 'means to an end' for Nolin in his investigation, the interest for the present study is that he, too, chose to evaluate student attitude using a 'verbal test' procedure.

One researcher to make further use of the assessment procedure employed by Nolin was Pogonowski who, in 1985, wished to determine "whether attitudes among students engaged in a Process-Oriented Music Curriculum (POMC) differ as a function of gender, grade level, socioeconomic status, and musical aptitude."²¹⁹ The Nolin inventory inspired

²¹⁶ Ibid.

²¹⁷ Ibid, p. 126.

²¹⁸ Ibid, p. 131.

²¹⁹ L. Pogonowski, "Attitude Assessment of Upper Elementary Students in a Process-oriented Music Curriculum", Journal of Research in Music Education, (1985), p. 248.

one of the instruments used by Pogonowski, as she chose to adopt the “I like it”, “It’s O.K.”, “I don’t like it”, and “I don’t remember doing it” response format.

A final study needing mention in this review of research is the work by Shaw and Tomcala, as reported in their article “A Music Attitude Scale For Use With Upper Elementary School Children”,²²⁰ published in 1976. The authors were motivated by “...the unavailability of a published, standardized music attitude scale suitable for use in upper elementary school grades ... to devise an attitude instrument and to begin the standardization process...”²²¹ Shaw and Tomcala chose the summated rating scale technique to develop their instrument; the final form being forty-four item statements to which the students could respond “on a four-point scale: strongly agree (4), agree (3), disagree (2), and strongly disagree (1).”²²² Shaw and Tomcala administered the scale to 97 students from a low socioeconomic inner-city minority school. “Instrument reliability was estimated to be .79 using the split-half method corrected by the Spearman-Brown formula. The factor analysis indicated there were 14 factors present...”²²³ [These factors were not further identified in the JRME report.] The scale was then “administered to 168 upper middle-class suburban children. Reliability for this administration was estimated to be .87 using the split-half method corrected by the Spearman-Brown formula. A factor analysis indicated there were 13 factors present.”²²⁴

Of the scale, the authors concluded that “...sufficient construct validity exists to warrant its use in studies for which another suitable instrument cannot be located.”²²⁵

²²⁰ C. Shaw and M. Tomcala, “A Music Attitude Scale for Use with Upper Elementary School Children”, Journal of Research in Music Education, (1976).

²²¹ Ibid, p. 75.

²²² Ibid.

²²³ Ibid, p. 76.

²²⁴ Ibid, p. 77.

²²⁵ Ibid, p. 80.

(2) Studies and research (into affective assessment in music) which are not directly related to the present topic:

This section - it will be recalled - will refer to literature and research which, although appearing to be related to the present study, contain aspects 'which preclude their direct applicability'. (p.72)

Working chronologically once more, the first study to identify is the "Tests of Musical Feeling and Musical Understanding" by Max Schoen²²⁶ - a report published in 1925 regarding three tests the author developed, those being 'Relative Pitch', 'Tonal Sequence', and 'Rhythm'. To return to Bullock's description of different types of musical attitude tests, the Schoen tests employ musical stimuli, therefore may be described as tonal (as opposed to verbal) tests. Schoen's interest in 'musical feeling' is more tied to musical 'appropriateness' than to attitude to music; and his 'musical understanding' draws on a subject's knowledge of musical concepts.

Soon to rise to prominence was the work of Kate Hevner, outlined in her 1930 and 1931 articles (as mentioned on p.73) At this stage of her research, Hevner was interested in developing and testing a diagnostic instrument with a view to measuring "...the appreciation of listeners for musical compositions played on the piano"²²⁷ (Hence not a verbal test). Nevertheless, of interest to the present study was the fact that Hevner was concerned with assessing affective domain elements; elements that hitherto were difficult to focus on objectively. In Hevner's words:

Esthetic qualities are not only elusive, but they are rather unstable, at least when present day measures of reliability are applied to them, and that

²²⁶ M. Schoen, "Tests of Musical Feeling and Musical Understanding", Journal of Comparative Psychology, (1925).

²²⁷ K. Hevner, (1930), op. cit., p. 470.

*“unerring good taste” which is so often claimed for certain individuals seems very difficult of demonstration.*²²⁸

Hevner's test was based on respondents discriminating between different versions of a musical composition: the original plus ‘spoiled’ versions. The assumption was that the original version of a composer's work was the aesthetically ‘correct’ one for a respondent to choose, that:

*...the creation of an artist whose work has been generally acclaimed for its merit is more beautiful than the same creation altered by a deliberate attempt to spoil its various beautiful qualities.”*²²⁹

The results of the Hevner ‘Music Discrimination Test’ were correlated with two noted other assessments, these being self-rating scales for musical training and musical talent respectively. Hevner's conclusion (perhaps somewhat optimistically) was that:

*...the preliminary study has shown that in spite of the increased difficulty encountered in presenting the musical test material to the subjects, it is nevertheless possible to employ this same technique and to develop a standardized objective test for appreciation of the beautiful in music which will attain a satisfactory degree of both reliability and usefulness.*²³⁰

In 1956, George Kyme's article “Are Musical Tastes Indicative of Musical Capacity?” appeared in the Journal of Research in Music Education.²³¹ Interestingly, the word ‘taste’ in the title is an example of how modern interpretations can generate misleading expectations: in this instance, Kyme proceeded to discuss his ‘Test of Esthetic Judgement’

²²⁸ Ibid.

²²⁹ K. Hevner, (1931), op. cit., p. 575.

²³⁰ Ibid.

²³¹ G. Kyme, “Are Musical Tastes Indicative of Musical Capacity?”, Journal of Research in Music Education, (1956).

(incidentally, a test employing musical stimuli), defining in his hypotheses the word 'taste' as being 'developed judgement'.²³² According to a 1986 'Glossary For Use in Affective Response Literature in Music' [developed using 'consultants'] the following definitions may be found:

- | | |
|-----------|--|
| judgement | - A critical evaluation or decision made after perception and discrimination. |
| taste | - A person's overall attitude toward collective musical phenomena. Long-term commitment to musical preferences. A social matter that tends to vary with varying groups of people, places, and times, and that gives the impression that preference for one kind of music is better than preference for another. ²³³ |

The Price definition of 'taste' obviously does not correspond to the context of the Kyme article - as 'attitude' was not a prime consideration. Hence an update of Kyme's title might read: "Is Esthetic Judgement Indicative of Musical Capacity?", particularly as Kyme himself appeared to favour that phrase.

Walter Lifton's 'Music Reaction Test', as described in "The Development of a Music Reaction Test to Measure Affective and Aesthetic Sensitivity"²³⁴, was a brave and innovative attempt at evaluating affective response to music. Lifton's rationale was, "In order to assess a subject's sensitivity to musical stimuli it was decided to present subjects with carefully chosen musical selections and solicit free written responses."²³⁵ The object of Lifton's research was to develop a hierarchy with which to classify the 'free' responses

²³² Ibid, p. 44.

²³³ H. Price, "Proposed Glossary for Use in Affective Response Literature in Music", Journal of Research in Music Education, (1986), p. 154.

²³⁴ W. Lifton, "The Development of a Music Reaction Test to Measure Affective and Aesthetic Sensitivity", Journal of Research in Music Education, (1961).

²³⁵ Ibid, p. 158.

engendered by the musical selections of the test; the resultant categories ranging from 1) fanciful thoughts and comments (most highly rated), to 5) Denials of feeling (least highly rated). One major weakness of the 'Music Reaction Test' not admitted by Lifton was that it presupposed that all respondents possessed the ability to write down their feelings and express themselves in prose.

A valuable study, however one without direct application to the present study due to its focus on a test with musical stimuli, is the PhD thesis of William Bullock (1971).²³⁶ In his thesis, Bullock outlined his "construction and evaluation of a test of musico-aesthetic attitude for use in comparing the effectiveness of various methods of music appreciation instruction at junior college level."²³⁷ 'Musico-aesthetic attitude' was defined by Bullock as being:

the composite of those opinions, feelings, interpretations, and judgments with which any individual identifies himself as he listens to a series of musical selections representing a variety of concert music styles. ²³⁸

He was primarily concerned with students' responses to concert music (as opposed to popular music of any kind). The developed test, the MAAP [Musico-Aesthetic Attitude Profile], consisted of 200 item statements describing ten pre-determined 'basic concert music styles'. Respondents were to accept or reject each of these in turn. Eleven scores were computed from each respondent's efforts; "one score on each of ten concert music style scales, and one composite score on the *Musico-Aesthetic Attitude Scale* (MAAS)."²³⁹ The reliability of the MAAS was calculated using the test-retest method, the result being .91. The ten style scales yielded reliabilities of between .57 and .86, with only two of these lying outside the .72 and .86 range. The undertaking of scale analyses "revealed that all

²³⁶ W. Bullock, Construction and Evaluation of a Test of Musico-Aesthetic Attitude, (1971).

²³⁷ W. Bullock, "Construction and Evaluation of a Test of Musico-Aesthetic Attitude", Dissertation Abstracts International, (1971), p. 5261 - A.

²³⁸ W. Bullock, Construction and Evaluation of a Test of Musico-Aesthetic Attitude, (1971), p.10.

²³⁹ W. Bullock, Dissertation Abstracts International, (1971), op. cit.

MAAP scales are capable of accurately ranking listeners.”²⁴⁰ A major drawback of the MAAP is that the instrument takes between 1 1/4 and 1 1/2 hours to complete (as reported by Bullock). However, an important feature of Bullock's test is that he recommends its use in “evaluating the level of musico-aesthetic attitude of individual students”;²⁴¹ many of the tests and instruments examined thus far are only effective in group assessments.

In a separate category from the studies described above, although worthy of mention in an examination of “music assessment research containing links with affective domain objectives”, is the research into musical aptitude testing undertaken from the 1930s by Herbert Wing.²⁴² The ‘Standardised Tests of Musical Intelligence’ developed by Wing consist of a battery of seven separate tests: three focussing on musical ability (Analysis of Chords, Pitch Discrimination, Memory for Pitch) and four measuring selected aspects of musical appreciation (Harmony, Intensity, Rhythm, and Phrasing Appreciation).

His musical appreciation tests may be interpreted as belonging to the affective domain, due to the fact that the respondent must exercise preference in choosing between differing versions of the same piece of music. The classification scheme adopted by Bullock, however, illustrates why the Wing tests, for the purposes of the present study, do not fall simply into a ‘tonal test’ category.²⁴³ [As discussed on p.72, “Verbal tests do not employ musical stimuli; tonal tests do. Verbal tests demand either retrospective or prophetic valuation because they do not use actual music.”²⁴⁴] According to Bullock, there is a further division of the ‘tonal test’ category - this being due to the distinction he makes between ‘evaluation’ and ‘valuation’ reactions:

²⁴⁰ Ibid.

²⁴¹ Ibid, p. 5262 - A.

²⁴² H. Wing, “Tests of Musical Ability and Appreciation”, The British Journal of Psychology Monograph Supplements, (1948).

²⁴³ Similarly with the “Musical Aptitude Profile” developed by Edwin Gordon: The three tests - Phrasing, Balance and Style - in his third category, “Musical Sensitivity”, might also be interpreted as belonging to affective domain assessment.

²⁴⁴ W. Bullock, (1973), op. cit., p. 336.

An evaluative reaction is a cognitive response involving critical appraisal of the stimulus based on perception of specific musical elements. A valuative reaction involves affective acceptance or rejection of the stimulus based on awareness of and attention to the stimulus, though not necessarily perception of its specific musical elements. ²⁴⁵

Bullock places the Wing Standardised Tests of Musical Intelligence into the 'evaluative reaction' division. As he explains, "A test subject's ability to perceive and/or evaluate specific aspects of the total musical stimulus rather than his opinion or evaluation of the stimulus is the variable in question."²⁴⁶

Of interest to this study, however, is Wing's attempt to assess the relationship between musical ability (as measured by his tests), and musical interest. In a specific study, Wing obtained self assessments concerning 'attitude towards music' from 333 adolescent schoolboys. "It was found that children had difficulty in assessing their own interest in music on any finely graded scale, and so they were asked to put themselves into one of four broad classes: A, very interested in music; B, interested; C, indifferent; and D, dislike."²⁴⁷

As a result of the survey it was calculated that approximately two-thirds of the boys had some interest in music (A=19%, B=48.5%), about a quarter were indifferent (C=26.5%), and 6% (D) actively disliked it.²⁴⁸

²⁴⁵ Ibid, p. 332.

²⁴⁶ Ibid, p. 333.

²⁴⁷ H. Wing, (1948), op. cit., p. 78.

²⁴⁸ Ibid, pp. 78-79.

Wing theorized that:

*Were ability to perform the tests derived from opportunities, the arousing of interest might be the first stage in its evolution; in that case a strong connection would be expected to exist between ability in the tests and interest.*²⁴⁹

Following the administration of the Wing Tests to the schoolboy group, a correlation coefficient between the interest and ability results was obtained: the value being 0.30. Wing concluded:

*The method of awarding marks to each grade [the method Wing used to calculate the correlation coefficient] must be arbitrary and cannot be expected to give more than a very approximate coefficient, but its smallness seems to indicate that ability is unlikely to have arisen as the result of interest. Questioning showed that interest was more likely to arise from ability, for, as was to be expected, children liked studying a subject in which success came easily, and disliked it if they experienced an uncomfortable feeling of inferiority.*²⁵⁰

3. STUDIES RELATED TO THE AFFECTIVE TAXONOMY OF KRATHWOHL, BLOOM AND MASIA.

The affective Taxonomy of Krathwohl *et al* has made an immense impact on researchers and educationalists, a number of whom have chosen to work with the model in order to further examine its applicability to evaluation and other educational requirements.

²⁴⁹ Ibid.

²⁵⁰ Ibid, p. 79.

(1) Use of the Taxonomy in general:

Given that Krathwohl's Taxonomy was published in 1964, it is of no surprise to discover a number of research studies, based exclusively on the model, appearing in the late 1960's.

Of direct interest to the present research is the PhD study of Arie Lewy, which he went on to report in The Journal of Experimental Education (1968).²⁵¹ Both the PhD thesis and the JEE article were entitled "The Empirical Validity of Major Properties of a Taxonomy of Affective Educational Objectives," and Lewy acknowledged the 'assistance and counsel'²⁵² of Professors B.B. Masia and B.S. Bloom.

In his work, Lewy sought to test the properties of the "Taxonomy of Educational Objectives: Affective Domain" of Krathwohl *et al*, thereby illustrating "... certain methodological problems involved in the general study of models".²⁵³ He began his JEE report by briefly outlining the basic structure of the Affective Taxonomy model, following with a 'requirements for validation' discussion. In this, Lewy attempted to "1) demonstrate the existence of empirical referents for its [the Taxonomy's] constructs, and 2) show that these empirical referents display the structure defined by the model."²⁵⁴ Lewy used as the sources of data for his validation research "... three special fields included in both high school and college curricula ... : Mathematics, Music, and Reading."²⁵⁵ He reported: "For each of these special fields, tests containing about 80 items at the Taxonomic levels of Receiving, Responding, Valuing, and Organization were constructed. Each of these was administered to from 200-300 high school and college students."²⁵⁶

²⁵¹ A. Lewy, (1968), op. cit.

²⁵² Ibid, p. 77.

²⁵³ Ibid, p. 70.

²⁵⁴ Ibid, p. 71.

²⁵⁵ Ibid, p. 72.

²⁵⁶ Ibid.

Lewy presented the preliminary drafts of the tests to “the judgement of experts in the three subject areas”²⁵⁷, and subsequently eliminated any test items thought to be inappropriate or otherwise unsatisfactory. This constituted the ‘Content Validity’ of the respective tests.

In order to obtain an estimation of the ‘Concurrent Validity’ of the tests, Lewy “...attempted to obtain evidence about relevant affective behaviors of the students on the basis of other types of responses which were independent of the tests to be validated.”²⁵⁸ In the case of the Music test, the ‘independent’ variable was the number of musical programmes the respondent had listened to during the week preceding the taking of the test. [This appears to be a weak point of Lewy's study, as it presupposes that all, or even most, students had access to radios or other music reproduction equipment which, in the 1960s, could not be assumed.]

With the formation and administering of the three tests, Lewy undertook to substantiate three hypotheses which he outlined in a Journal of Research in Music Education article:

- (1) Consensus can be obtained among expert judges as to the proper taxonomic level of given affective behavior.
- (2) In any given population, behaviors at a higher taxonomic level will occur less frequently than those at a lower level.
- (3) The degree of association between behavior at adjacent taxonomic levels is higher than the degree of association between behaviors at nonadjacent levels.²⁵⁹

In support of the first hypothesis, Lewy submitted all test items to two experienced ‘raters’, who were instructed to independently classify each item “...according to the categories of

²⁵⁷ Ibid.

²⁵⁸ Ibid.

²⁵⁹ A. Lewy, “Affective Outcomes of Musical Education”, Journal of Research in Music Education, (1971), p. 362.

the model.”²⁶⁰ The results were markedly positive: the raters “...agreed on the allocation of 74 percent of the items, and assigned 25 percent of the remaining items to adjacent, and only one percent to non-adjacent categories.”²⁶¹ Lewy concluded that such a level of agreement:

*...indicates that the Taxonomy can be used as a device for classifying affective educational objectives and outcomes.”*²⁶²

Following the administration of the three tests, Lewy was able to examine and interpret data with reference to the second and third hypotheses. In the first instance, he wished to establish that “the frequency with which a given behavior occurs will vary inversely with its taxonomic level.”²⁶³

This Lewy was able to accomplish, as “...in each test the percentage of students displaying a given behavior decreased from lower to higher level taxonomic categories.”²⁶⁴ There was, however, one exception, that occurring in the music test “... where behaviors at level 4 (Organization) were more commonly exhibited than were those at level 3 (Valuing).”²⁶⁵ [Lewy’s comment at this point was “The deviation from the hypothesized structure of items in level 4 may suggest that measurement of affective behavior at that level calls for much refinement and improvement.”²⁶⁶] Finally, Lewy addressed the third hypothesis; in other words, that “the degree of positive correlation between any two variables is inversely related to the hierarchical distance by which they are separated.”²⁶⁷ This he was able to support, after undertaking various statistical analyses.

²⁶⁰ A. Lewy, (1968), op. cit., p. 73.

²⁶¹ Ibid.

²⁶² Ibid.

²⁶³ Ibid.

²⁶⁴ Ibid, pp. 73-74.

²⁶⁵ Ibid, p. 74.

²⁶⁶ A. Lewy, (1971), op. cit., p.364.

²⁶⁷ A. Lewy, (1968), op. cit., p. 73.

Lewy thus concluded that:

*...the constructs of the model have empirical referents among affective educational objectives, and that the hierarchical structure of these referents corresponds to that claimed by the model.*²⁶⁸

A second doctoral study to make use of the affective Taxonomy was undertaken in 1968 by John Leslie Purcell. His Ed.D. thesis, entitled Developing an Attitude Scale Using the Taxonomy of Educational Objectives - Affective Domain, examined "...the possibility of using the rationale and continuum presented by Krathwohl, Bloom, and Masia as an aid in developing and ordering items for acceptable attitude scales".²⁶⁹ The argument was based on the idea that:

*while the Taxonomy did not offer new definitions, it did provide a conceptual model which makes it easier to visualize the relationships between affective variables and also made it easier to define the variables in terms of their behavioral manifestations.*²⁷⁰

Purcell was of the opinion that traditional ways of constructing attitude scales, namely the Likert and Thurstone methods, were both "time consuming and technically involved"²⁷¹, and he believed that the Taxonomy of Krathwohl *et al* could provide the foundation for a more practical alternative. In order to examine the effectiveness of the Taxonomy rationale and continuum Purcell presented four questions as follows:

- 1. If attitude statements are written to coincide with various levels of the Taxonomy continuum, will they be properly ordered to indicate the degree to which an attitude is held?*
- 2. Will scales constructed from these statements prove to be reliable?*

²⁶⁸ Ibid, p. 76.

²⁶⁹ J. Purcell, "Developing an Attitude Scale Using the Taxonomy of Educational Objectives - Affective Domain", Dissertation Abstracts International, (1968), p. 519 - A.

²⁷⁰ J. Purcell, "Developing an Attitude Scale Using the Taxonomy of Educational Objectives - Affective Domain", Unpublished Ed.D. Thesis, Washington State University, (1968), p.5.

²⁷¹ Ibid, pp. 5-6.

3. *Will such scales give evidence of unidimensionality?*

4. *Will such scales discriminate between subjects or groups of subjects in a meaningful way?* ²⁷²

Six hypotheses were prepared, based on the above questions. The first two involved comparing the 'ordering' of attitude scale items according to the Taxonomy continuum with the independent ordering ("by a group of 25 judges"²⁷³ [see footnote, p.73 for reference to the 'judging group' technique]) of the same items according to 1) the guide of 'favourable' to 'unfavourable', and 2) the principle of equal-appearing intervals. The scales developed were designed to measure 'Willingness to Use Democratic Classroom Procedures' and 'Attitude Toward Racial Integration'.

*Comparisons were made, using the Spearman rank-order correlation method, between the scale order established by a group of 25 judges, and the order established using equal-appearing interval scale values.*²⁷⁴

Purcell reported that the lowest correlation coefficient obtained was .75, with most of the values exceeding .80. He described these results as being 'good', and subsequently concluded that:

*...the order imposed on the items in the scales by the Taxonomy continuum was essentially the same as the order established by the other methods, and that the levels of the Taxonomy continuum are correctly ordered.*²⁷⁵

Purcell administered the resulting scales (six in all ²⁷⁶) to 150 students, with a view to testing four further hypotheses, as follows:

²⁷² Ibid, p. 10.

²⁷³ J. Purcell, Dissertation Abstracts International, (1968), op. cit.

²⁷⁴ J. Purcell, Ed.D. Thesis, (1968), op. cit., p. 37.

²⁷⁵ J. Purcell, Dissertation Abstracts International, (1968), op. cit.

²⁷⁶ Scales I and II, and Scales IV and V, were ordered according to the Taxonomy continuum, and pertained to the Willingness to Use Democratic Classroom Procedures and Attitude Toward Racial Integration respectively. Similarly, Scale III and Scale VI were ordered using the equal-appearing interval method.

Hypothesis III: Scales constructed from items developed using the Taxonomy rationale will be found to be reliable using stability of measurement as a criterion.

Hypothesis IV: Scalogram analysis of the results of the administration of the scales ... will provide evidence of unidimensionality.

Hypothesis V: There will be no significant difference between the scores of women students and the men students who serve as subjects for this study.

Hypothesis VI: There will be no significant difference between the scores of the subjects tested who were preparing to teach at the elementary level and those who were preparing to teach at the secondary level.²⁷⁷

His method of scoring became relatively convoluted however, as he proceeded to calculate seven separate scores for the four Taxonomy-developed scales, and three scores for the remaining three scales. In order to calculate the various reliability figures (for Hypothesis III), Purcell chose the test-retest method, and used the Pearson Product-Moment Correlation method to compute the coefficients. The results he received, as he himself admitted, “varied greatly with the method of scoring and the method of establishing the scale order.”²⁷⁸ He found that the best results were achieved when “the total scores were combined for Scales I and II and for Scales IV and V, [when] reliability coefficients of .78 and .89 were obtained.”²⁷⁹ Purcell considered these coefficients to be ‘quite satisfactory’, and concluded that:

*the scales constructed for this study from items developed using the Taxonomy rationale were reliable using stability of measurement as a criterion and that the data supported the hypothesis as stated.*²⁸⁰

²⁷⁷ J. Purcell, Ed.D. Thesis, (1968), op. cit., p. 10.

²⁷⁸ Ibid, p. 29.

²⁷⁹ Ibid, p. 39.

²⁸⁰ Ibid.

As far as Hypothesis IV was concerned, scalogram analysis of the results obtained yielded coefficients of reproducibility of between .84 and .92. Purcell felt that these “appeared to support the conclusion that the scales were unidimensional.”²⁸¹ However, he cautioned that more evidence was required - perhaps from administering the scales to other groups - before making a ‘final judgement’ on this matter.

Hypotheses V and VI were found to be supported, following comparisons of the appropriate mean scores.

Purcell concluded that:

*the Taxonomy rationale and continuum can be effectively used as an aid in developing and ordering items for acceptable attitude scales.*²⁸²

Despite undertaking a valuable investigation, Purcell appeared to be somewhat less than rigorous in recounting his research efforts. Of particular concern is his lack of detail regarding the 150 subjects of the testing group, who are referred to only as “students in the Educational Psychology classes at Washington State University” (pp. 22 and 32 of the Thesis); and the fact that the test-retest results were calculated from administering the scales to only 42 students (also from the Educational Psychology classes). Nevertheless, Purcell's work at least contributed to a growing pool of studies supporting the use of the ‘Taxonomy’ in affective domain assessment.

A third study focussing on the affective domain Taxonomy was undertaken in 1969 by Eiss and Harbeck. Working under the umbrella of the National Science Supervisors Association, the authors published a booklet (Behavioral Objectives in the Affective Domain)²⁸³, in which they examined the affective domain as it relates to science education.

²⁸¹ Ibid, p. 40.

²⁸² Ibid, p. 41.

²⁸³ A. Eiss and M. Harbeck, Behavioral Objectives in the Affective Domain, (1969).

Although somewhat condensed, Behavioral Objectives in the Affective Domain is a reasonably comprehensive document, as Eiss and Harbeck spent some time discussing such factors as (1) the affective domain as it relates to both the cognitive and psychomotor domains, (2) using behavioral objectives in evaluation, (3) the various components of the affective domain - as set down by Krathwohl *et al* - including attitudes, interests, values and motivation, and (4) procedures for evaluating affective outcomes, with a descriptive list of different types of affective evaluation instruments - for example, interviews, inventories, and rating scales. Although intended for use in science education, this booklet contains information and suggestions which are widely applicable to any field of education concerned with formalising an approach to behavioral assessment in the affective domain.

In the same year, a group of educationalists saw the need to work further with both the cognitive and affective domain Taxonomies as far as their direct applicability to the school setting was concerned. Metfessel, Michael, and Kirsner, set out their proposals in an article entitled "Instrumentation of Bloom's and Krathwohl's Taxonomies for the Writing of Educational Objectives".²⁸⁴ The authors intended to develop "...behaviorally oriented infinitives which, when combined with given objects, would form a basis for meaningful, cohesive, and operational statements."²⁸⁵ Their work resulted in the construction of tables: these were a collection of key words for 'infinitives' and 'direct objects' for each level of the Taxonomies - from which it would be possible to write appropriate objective statements. For example, for the first level of the affective Taxonomy ('Receiving'), the '1.1 - Awareness' sublevel had listed as its 'infinitives': to differentiate, to separate, to set apart, to share; while its corresponding 'direct objects' were: sights, sounds, events, designs, arrangements. At the opposite end of the continuum, the 'infinitives' for the 5.2 sublevel of 'Characterization' were: to be rated high by peers in, to be rated high by

²⁸⁴ N. Metfessel, W. Michael and D. Kirsner, "Instrumentation of Bloom's and Krathwohl's Taxonomies for the Writing of Educational Objectives", Psychology in the Schools, (1969).

²⁸⁵ *Ibid*, p. 227.

superiors in, to be rated high by subordinates in; with the 'direct objects' being: humanitarianism, ethics, integrity, maturity.

The authors reported that:

*these tables have been of considerable help to ... students, as well as to personnel in public schools who are concerned with writing objectives prior to curriculum development, constructing test items, or to carrying out evaluation studies.*²⁸⁶

Also concerned in 1969 with the general application to education of the cognitive and affective Taxonomies was R.H.Dave, who, in his article "Taxonomy of Educational Objectives and Achievement Testing" sought to "...derive some major implications of taxonomies to achievement testing."²⁸⁷ Dave took a more general approach to the application of the Taxonomies than did Metfessel *et al* (who were concerned with the precise details of constructing objectives statements); he was more concerned with the underlying rationale. Dave identified (and briefly examined) the benefits of basing an evaluation programme on the Taxonomies as being the following:

- (a) Building logic in testing.
- (b) Making evaluation more comprehensive.
- (c) Criterion for evaluating an evaluation device.
- (d) Developing a new typology of test items.
- (e) Aid to constructing objective-based test items.
- (f) Aid to planning a test and interpreting results.

²⁸⁶ Ibid, p. 230.

²⁸⁷ R. Dave, "Taxonomy of Educational Objectives and Achievement Testing" in Developments in Educational Testing, Volume 1, ed: K. Ingenkamp, (1969), p. 211.

Both Purcell and Dave drew attention to the fact that using the Taxonomies for evaluation purposes requires certain expertise and experience: Purcell suggested that one would need to investigate "...whether prospective teachers and others can be taught, expediently, to use the [affective] Taxonomy efficiently for the purpose of attitude appraisal."²⁸⁸ He made the reminder however, that

*...it should be noted that the writer and two judges, working independently and without consultation, did reach close agreement on 68% of the original items. This agreement resulted from one attempt to classify the items, and none of them were rewritten as a result of the comparison of the three judgments. This would seem to imply that the rationale can be understood and applied without a great deal of ambiguity.*²⁸⁹

Dave identified as difficulties the construction of test items for the higher levels of the Taxonomies, and distinguishing between consecutive categories and sub-categories.

However, he suggests:

*With sufficient practice, experience and familiarity with a large variety of objective-based test items the evaluator acquires greater confidence in using the taxonomy.*²⁹⁰

In addition to the above-mentioned studies and research, further references to the Taxonomies can be found in the subsequent writings of Bloom *et al*, namely, in Handbook on Formative and Summative Evaluation of Student Learning, and in Evaluation to Improve Learning.

(2) Use of the Taxonomy and Music:

It is necessary at this point to return to the PhD study of Arie Lewy, as the inclusion of a music test in his research constitutes one of the earliest links between the Krathwohl

²⁸⁸ J. Purcell, Ed.D. Thesis, (1968), op. cit., p. 41.

²⁸⁹ Ibid.

²⁹⁰ R. Dave, (1969), op. cit., p. 214.

Taxonomy and assessment in music. Despite the fact that the test was only a minor part of his investigation, Lewy realised the ramifications and wrote a separate article three years later, for the Journal of Research in Music Education. In this article "Affective Outcomes of Musical Education"²⁹¹, Lewy isolated the Music component of his doctoral study, relating it, with implications, to the field of Music Education. His suggestions included "(1) that the categories of the affective taxonomy are suitable in describing different types of affective behavior toward music, and (2) that these types of behaviors can be arranged on a hierarchical scale."²⁹² He outlined the benefits of utilising the affective taxonomy, particularly as far as formulating educational objectives is concerned. He concluded:

*It is to be hoped that measurement instruments constructed in terms of the taxonomy will provide information about affective outcomes in the field of teaching music and will contribute to the improvement of teaching procedures.*²⁹³

Lewy did not undertake any further work with his music assessment instrument.²⁹⁴

The first established music educator to direct attention to the affective Taxonomy of Krathwohl *et al* was Richard Colwell who, in 1970, devoted an entire appendix section of The Evaluation of Music Teaching and Learning²⁹⁵ to its interpretation and application. This was a systematic attempt to relate the different categories and sub-categories of the taxonomy to the teaching and evaluation of general music. The examples provided at each level were specific and detailed, giving some ideas, for example, as to the item statements that attitude or preference tests might contain. Despite the apparent detail, however, the assessment formats suggested at each level were not consistent: Colwell may not have had a specific testing instrument in mind. He gave no indication of scoring procedures.

²⁹¹ A. Lewy, (1971), op. cit.

²⁹² Ibid, p. 365.

²⁹³ Ibid.

²⁹⁴ This was indicated in a letter from Lewy to the author.

²⁹⁵ R. Colwell, (1970), op. cit., pp173-177.

Colwell also presented an overview of the affective domain, and of various assessment measures in the chapter “evaluating the affective domain”. Ways of assessing affective objectives were examined historically (although briefly) - including Thurstone’s and Likert’s methods of attitude scale construction. Comments were made under the respective headings of ‘Rating Scales’, ‘Attitude Scales and Projective Tests’, and ‘Questionnaires, Checklists, and Other Devices’,

Altogether, Colwell's contribution to the affective domain and music education was particularly useful, as he attempted to isolate and discuss aspects of affective domain assessment in more detail than was common in other texts of the time (see pp.2 - 6).

In 1976, Robert Noble published the results of his research regarding “A Multivariate Analysis of Factors in the Backgrounds of Wyoming Adults Related To Their Attitudinal Levels Concerning Music”.²⁹⁶ As outlined in the report of that study, “Attitudinal level was defined as any one or more of a possible 11 achieved hierarchical levels of response, adjudged as appropriate by a jury, to an instrument designed to measure attitudes toward music”.²⁹⁷

Of immediate interest to a study of assessment in the affective domain is Noble's method of measuring attitudes towards music. The instrument he devised, an inventory, was

*based on the work of Krathwohl and others as analyzed for music by Colwell. (The 5.0 level, “Characterization by a value or value complex” was omitted from the inventory).*²⁹⁸

Noble referred to Colwell's The Evaluation of Music Teaching and Learning, pp.173-177, as a source for his inventory development, but did not supply particular details. One would assume that he followed the guidelines set down by Colwell in his Appendix section: see p. 99 of the current study for reference to this work.

²⁹⁶ R. Noble, “A Multivariate Analysis of Factors in the Backgrounds of Wyoming Adults Related to their Attitudinal Levels Concerning Music”, Bulletin of the Council for Research in Music Education, (1976), also in A Cross-Section of Research in Music Education, ed: S. Barnes, (1982).

²⁹⁷ R. Noble, (1982), op. cit., p. 230.

²⁹⁸ Ibid, p. 232.

With regard to Noble omitting the 5.0 taxonomic level, one has to make the connection with an earlier statement he made, viz: “the investigator sought to measure attitudes only toward music (a portion, only, of the affective domain).”²⁹⁹ Noble didn’t specify the position of ‘attitude’ on the Krathwohl affective continuum: as will be recalled from p.39, attitude spans the taxonomic levels 2.2 - 4.1. Presumably he subscribed to this theory, however.

Noble’s eleven hierarchical levels of response, then, corresponded to the sublevels of 1.1 - 4.2 of the affective taxonomy. [Refer to Appendix D of this thesis for exact details.]

Noble began with a 126-item instrument which was presented to 149 Wyoming adults for item testing, and to 18 music education specialists for suggestions. Factor analyses followed, resulting in the retention of 57 items, each representative of one level of the taxonomy of Krathwohl *et al.* The theory behind the inventory was thus: “In the first ten levels of response, if the respondents’ opinions agreed with those the jury perceived as ‘appropriate responses’ to four out of five inventory items measuring that level, the respondent was adjudged to have ‘met’ that level. In the highest level, level 11, the respondent had to be in agreement with jury judgement in five out of the seven inventory items to have ‘met’ that level. Inventory items were not grouped together by hierarchical level; rather, they were rearranged to avoid patterns in either questions or responses.”³⁰⁰

Noble proceeded to use his developed instrument to test specific hypotheses of his research, these last stemming from Noble's main question: “What is the relationship between the quantitative amount of music education and attitudes toward music of Wyoming adults?”³⁰¹ The instrument itself was not tested further, a limitation that Noble himself admitted in the conclusion of his report: “Since validity and reliability of a criterion instrument are critical to the integrity of conclusions, the latter may be somewhat suspect in

²⁹⁹ Ibid.

³⁰⁰ Ibid, p. 233.

³⁰¹ Ibid, p. 229.

this research because factor analysis only was employed as a basis for inclusion of inventory items.”³⁰² Furthermore, Noble chose not to question the basic principles of the Krathwohl Taxonomy as a model on which to base his assessment instrument, stating: “The levels of attitudinal response theoretically exist in this research on a hierarchical continuum.”³⁰³ Although he did not refer directly to the work of Lewy (his JRME article was quoted in the references section), Noble would have been accepting his research results with regard to the existence in the taxonomic model of a hierarchy pattern.

4. SUMMARY

This review of the research into affective domain assessment - with particular reference to music education - has illustrated the on-going concern of educationalists to further ‘demystify’ the evaluation of such aspects in music as attitudes, interests, feelings and appreciation.

The intention of the review was to provide a background for the present investigation; essentially establishing a need for further inquiry. Almost without exception, the studies cited have taken as their point of departure the fact that not enough research has been undertaken into the development of affective domain assessment instruments (for example, Hevner (1930), Long, Shaw and Tomcala, Bullock (1971), and Whybrew (1973)), a situation that they seek to redress. [Despite, however, taking a structured and systematic approach, a number conclude by admitting problems with establishing validity and reliability (for example, Farnsworth, Edwards and Edwards, and Shaw and Tomcala).]

The underlying criterion for the studies and research included in this Review was the fact that they are concerned with issues closely allied to the present investigation. The need arose, however, to supplement the discussion with the section dealing with “literature and

³⁰² Ibid, pp. 244-245.

³⁰³ Ibid, p. 233.

research which, although appearing to be related, contain aspects which preclude their direct applicability.” This section was included “so as to further delimit the topic in hand.” (p.72)

Predominantly then, the Review sought to follow the lead set by Kate Hevner's ‘Tests For Attitude Toward Music’, with regard to the verbal testing of affective elements in music. In other words, assessment procedures that are reliant on direct musical stimuli, and music aptitude testing, were not automatically covered as part of the main discussion.

The affective Taxonomy of Krathwohl, Bloom and Masia received prime attention in this review, as it is a recommendation of the current study that the Taxonomy constitutes a valuable educational model on which to base an affective assessment instrument. As demonstrated by Lewy, Purcell, and Noble, it is quite feasible to construct an assessment instrument consisting of item statements representing the various levels of the Taxonomy. The precedent set by these researchers is one which is highlighted in the current investigation, particularly in the quest to meet “the challenges associated with the affective domain and music education.” (p.40)

PART TWO

CHAPTER V - A NEW MUSIC ATTITUDE ASSESSMENT INSTRUMENT

The proposal of the current study [as introduced in Chapter II] was to develop a music attitude assessment instrument for use with New Zealand adolescent-age school students, based on the Taxonomy of Krathwohl *et al.*. To reiterate, the word ‘music’ here embraces all genres (for example, from popular to classical styles) and takes into account any structured or casual involvement in musical pursuits (including community, school or ethnic activities) [from p.41].

1. PILOT STUDY

As suggested by Arie Lewy, however, one must firstly “...raise questions as to whether [a] proposed model can be of practical use in dealing with issues related to the field of education.”³⁰⁴ Lewy identifies two such questions, and these will, in fact, assist in shaping the present research. The first of his questions asks if “...affective behavior towards music is classifiable according to the taxonomy categories”; the second, whether “...such behavior exhibits a hierarchical structure as delineated by the model.”³⁰⁵

In this investigation, the second question is addressed first; in other words, effort is made from the outset to establish whether a hierarchical pattern results from administering an ‘affective assessment’ instrument, which has been based on the Krathwohl Taxonomy. If this pattern does not emerge, one could suspect one or more of the following three factors:

- (1) The ‘hierarchical’ element of the Taxonomy of Krathwohl *et al* does not exist.
- (2) Affective behaviors are not able to be classified hierarchically.

³⁰⁴ A. Lewy, “Affective Outcomes of Musical Education”, Journal of Research in Music Education, (1971), p. 362.

³⁰⁵ Ibid.

- (3) Affective behaviors do not exhibit a hierarchical structure which will correspond to that of the Krathwohl Taxonomy.

According to Lewy, one would wish to discover that “In any given population, behaviors at a higher taxonomic level will occur less frequently than those at a lower level.”³⁰⁶ In order to pursue this line of inquiry, it was decided to readminister his affective assessment (in music) instrument based on the Taxonomy of Krathwohl *et al* ³⁰⁷, as it had already been used ‘successfully’ in past research. Due to the fact that this instrument was designed some years ago and needed modifying (see below for details), one could not expect a ‘perfect’ outcome in view of the changed circumstances. An ‘indication of’, or ‘tendency toward’ a hierarchical pattern in the results would be all that one could reasonably expect.

For the purposes of the current study the aim was to neutralise, as far as possible, any specific orientation of the Instrument that would disadvantage respondents either culturally or socially. For example:

1. Some items identified a particular musical genre: this would alienate students who had limited musical experiences, e.g. “I do enjoy modern symphonic music”, and “To me there are few things which could be more boresome like grand operas and symphony concerts.”
2. A number of items included the word ‘concert’: in the context of this instrument the concept of ‘going to a concert’ is restricting and culturally biased. It could be acceptable if it is defined in broad terms, e.g. classical concert *or* jazz concert, school music concert, community concert, rock concert etc. Hence when modifying such items, the phrase “or other live musical performance” is included.
3. Some items incorporated outmoded language, e.g. “Most types of music are effeminate and have little possible interest for persons engaged in more active pursuits of life”.

³⁰⁶ Ibid.

³⁰⁷ As developed in his PhD study, The Empirical Validity of Major Properties of Affective Objectives.

4. Some items were unsuitable for a general, unselected adolescent-age respondent group, as they required more sophisticated value judgements³⁰⁸, e.g.:

“Many literary themes have musical and drama versions, e.g. Faust, etc. Have you ever tried to compare the merits and demerits of the drama version and musical version of the same theme?”,

and

“The aesthetic perfection of truly great music surpasses that of all other artistic forms”.

In some cases a small modification of items was all that was required, in others none was possible, and the items had to be dropped from the inventory. [see Figure 1, p.107]

An example of an item requiring modification was:

Coming across familiar melodies when listening to music causes me great pleasure. [A level two statement]

This was modified (due to the term ‘melodies’ 1.being too restrictive, and 2. not being a common one in adolescent vernacular) to become:

Coming across familiar music - songs etc - causes me great pleasure.

An example of an item unsuitable for modifying was:

To me there are few things which could be more boresome like grand operas and symphony concerts. [A level two statement]

This was unsuitable for modifying due to the fact that it is culturally skewed: it presupposes knowledge of ‘grand opera, and symphony concerts. Furthermore, there is a negative implication in the use of the word ‘boresome’ that could sway some respondents’ answers.

³⁰⁸ - than all members of an unselected group would be capable of making.

The procedure for the modification of items was:

- (1) identify the specified taxonomic level of the item (see below, ‘item categories’).
- (2) adjust the item according to the requirements of that level, using both Krathwohl and Bloom’s descriptions, and the suggestions outlined by Colwell, in The Evaluation of Music Teaching and Learning.³⁰⁹

The complex and specific nature of the items of the higher levels precluded modification.

Figure 1: The adapting of the Lewy Instrument [Refer Appendices E-H]

Item Categories:

LEVEL	ITEMS
1	1 2 3 4 5 8 21 34 42 47 48 53 54 61 67 68 69 70 71
2	7 9 10 12 16 23 24 25 27 29 30 32 36 72 73 78
3	6 11 13 14 15 17 18 19 20 22 26 31 33 37 40 41 51 52 55
	56 57 58 59 60 62 63 64 65 66 74 75 76 77
4	28 35 38 39 43 44 45 46 49 50

Suitable: 1 2 5 15 19 20 21 22 26 28 30 31 33 34 39 44 45 46 50 55 56 57
61 63 65 66 67 69 71 72 73 77

Needed modifying: 4 6 7 8 9 10 11 12 14 16 23 27 32 35 36 38 43 51 64 74
75 78

Unsuitable for modifying: 3 13 17 18 24 25 29 37 40 41 42 47 48 49 52 53 54
58 59 60 62 68 70 76 79-86

³⁰⁹ R. Colwell, The Evaluation of Music Teaching and Learning, (1970).

Any modifications required were slight, and no substitutions were made. The percentage of items retained from the original Instrument was 63%, with items representing a spread of four levels of the Krathwohl Taxonomy.

This adapted version of the Lewy Instrument [see Appendix G] was administered to a group of ten third-form general music students [of mixed academic abilities] - to detect initially any specific problems inherent in either the basic format or the language used. The students coped with the instrument well, and there were no reported problems associated with the overall design. It was suggested by the teacher involved that students in today’s New Zealand schools are very much familiar and comfortable with the multi-choice format. As a result, only two changes were necessary. [Items 15 and 27 contained a double negative - “It is not unusual...” - which was confusing for the respondents. This was amended to “It is quite usual...”, and “It is usual...”] This was the desired outcome, as the aim was to keep any adjustments to a minimum, thus holding the adapted version as close to the original as possible.

To score, positive responses to the items deemed to belong to Level One were totalled, followed by those of Level Two, Three and Four. Thus each student was accorded four ‘scores’, which were expressed as percentages to allow for some comparability between the represented Levels of the Taxonomy hierarchy. (In the adapted version, the total number of items retained at each level varied considerably). Statistically, the results of this first trial were as follows:

Table 1: Results of Adapted Lewy Instrument - 1st trial

A	B	C
4	6	0

where:

- A = 'Scores' gained at each taxonomic level follow a distinct hierarchy pattern (for example, '80%-60%-40%-30%').
- B = Slight deviation from hierarchy pattern, i.e. 'scores' from three of the four taxonomic levels represented follow a hierarchy pattern (for example, '70%-80%-40%-30%').
- C = There is more than one disruption in the hierarchical sequence (for example, '60%-80%-30%-40%').

In other words, the 'four score' sequences of all ten students followed exactly, or had only a slight deviation from, a hierarchy pattern.

It will be recalled from page 105 that the aim was to identify "...an 'indication of', or 'tendency toward' a hierarchical pattern..." in the trialling of the adapted affective assessment Instrument, and to that end, the results reported above were considered supportive.

Other benefits emerged from this initial project, mainly relating to the practicalities of developing an instrument of this nature. In particular, the issue of semantics was raised, and the need to test thoroughly the level of language adopted for New Zealand adolescent-age subjects.

The adapted Lewy Instrument was therefore administered to two further groups of secondary school general music students, those groups being (1) 40 fourth form students (14-15 year olds), and (2) 25 third and fourth form students from another city. The results were as follows (A, B, and C as above):

Table 2: Results of Adapted Lewy Instrument: 2nd trial

	No.	A	B	C
Group 1	40	10%	60%	30%
Group 2	25	28%	40%	32%
Together	65	17%	52%	31%
Total Trialled	75	20%	53%	27%

From Table 2 it can be seen that, out of a total of 75 thirteen to fifteen year olds, 73 percent followed exactly, or were close to (i.e. ‘showed a tendency toward’), a hierarchy pattern. 27 percent showed no pattern at all. Although some differences were recorded with the weightings of categories A and B, the percentages in category C were similar for all groups in the second trial.

As these results indicate the possibility that behavior in the affective domain “...exhibits a hierarchical structure”, it was concluded that there was a strong case for basing an affective assessment Instrument on the Taxonomy of Krathwohl *et al.*

A second question was posed by Lewy [refer p.104] with regard to examining whether a “...proposed model can be of practical use in dealing with issues related to the field of education”.³¹⁰ He wished to discover if “...affective behavior towards music is classifiable according to the taxonomy categories”, and the procedure he adopted is outlined on pages 89-92. The hypothesis he formed - that “Consensus can be obtained among expert judges as to the proper taxonomic level of given behavior”³¹¹ - was re-examined in the light of the present study.

³¹⁰ A. Lewy, (1971), op. cit., p. 362.
³¹¹ Ibid.

Two educationalists were each submitted a copy of the adapted Lewy Instrument, along with a descriptive outline of the first four Levels³¹² of the Krathwohl Taxonomy. Their instructions included the following statement:

“What I ask of you now is that you consider each ‘item’ in turn, identify in your own mind the objective you feel is being represented, and nominate the appropriate taxonomic category (i.e. one to four). There is no one ‘correct’ outcome to this exercise, nor need there be equal numbers of statements/questions in the various levels.”

The responses of the educationalists were compared with Lewy’s placement of items, and the outcome was analysed according to the procedure adopted by Purcell in his research (refer Chapter IV). Purcell explained:

If three persons, the writer and the two judges, could agree on the classification of the item, it could be assumed that it was correctly written to correspond to the particular Taxonomy level. For the purposes of this study, it was decided that agreement would consist of having two of the three (writer and two judges) agree as to the exact subcategory placement, provided that the third person’s classification did not vary more than one subcategory above or below. ³¹³

In the Purcell study, it was reported that “unanimous or two out of three agreement was reached for a total of 68 items in each master scale.” ³¹⁴ As there were 100 items in each master scale, the consensus percentage was 68%. This compared with 69% achieved in the present study. Purcell felt that the 68% represented “considerable agreement” ³¹⁵, thus

³¹² It will be recalled that the Lewy instrument only used the first four levels of the Krathwohl Taxonomy.

³¹³ J. Purcell, (1968), op. cit., p.14.

³¹⁴ Ibid, p.24.

³¹⁵ Ibid.

indicating that, even without subsequent discussion with the judges, acceptable consensus could be reached as to the taxonomic level of items.

In short, the results that were obtained in all aspects of the readministering of the Lewy Instrument were certainly positive, thereby adding support to the conclusions Lewy himself made.

2. DEVELOPMENT OF THE NMAAI

Having established support for utilising the Taxonomy of Krathwohl *et al* in the field of affective assessment, it was now time to develop a new version of a music attitude assessment Instrument, one which was not tied to the requirements of past research. The criteria for this new Instrument were that the contemporary issues of cultural and social unfairness, individuality, and formative evaluation would be addressed. [From list, p. 51]

One may recall the statement by Shaw and Wright regarding 'attitude' research, that "...contributions might have been greater if research had been more cumulative in nature."³¹⁶ Also that "...the quality of the measuring instruments often is poorer than it would have been if existing scales had been used and improved."³¹⁷

It was therefore decided to build on the work already undertaken in the pilot study and select many of the item statements from prior research. A collection of item statements was duly compiled, the sources being the adapted Lewy Instrument, the Noble Instrument [see Appendix C], and the 'Tests for Attitude Toward Music' of Kate Hevner [see Appendix B], with the addition of some originals. The item statements were chosen particularly for their appropriateness and relevance to today's students. With respect to the last criteria, 'appropriateness' and 'relevance', one must make mention at this point the changing

³¹⁶ M. Shaw and J. Wright, Scales for the Measurement of Attitudes, (1967), p. ix.

³¹⁷ Ibid.

attitudes of contemporary music teachers toward the various music genres. The original Lewy Instrument, for instance, was developed to ascertain student attitude to 'classical' music only. Traditionally, 'good' music tended to consist solely of established 'art' music: popular and jazz idioms were automatically considered to be unsavoury. (In 1941 the New Zealand music educator, Vernon Griffiths, for example, felt that teenagers' inherent immaturity needed to be counteracted for their own benefit. His teaching method included drilling the students with 'good' music in order to lead them "from the primitive conceptions of inexperience to the more balanced judgement and clearer perception of maturity."³¹⁸ Reporting on his results, Griffiths proudly stated that "after a few weeks of daily assembly, at which jazz and swing music were unknown and folk songs and national songs took their place, the boys and girls showed clearly that they had a healthy musical appetite.")³¹⁹

As far as music education in today's New Zealand secondary schools is concerned, a more flexible approach needs to be adopted. Hence the attitude assessment Instrument being developed in the present study recognises the differing musical preferences of the students, thereby incorporating a more 'universal' perspective.

Due to their differing origins, some of the item statements at this stage required yes/no answers, others true/not true replies. The 'true/not true' format was eventually to be adopted as the most suitable. [See below for further discussion]

There was a sum total of forty item statements, which in turn broke into four groups of ten, each corresponding with one of the first four levels of the Krathwohl Taxonomy. The exclusion of level five in the constitution of the Instrument was a principle observed by both Lewy and Noble. As Krathwohl himself stated:

³¹⁸ T.V. Griffiths, *An Experiment in School Music Making*, (1941), p. 74.

³¹⁹ Ibid.

Rarely, if ever, are the sights of educational objectives set to this level of the Affective Taxonomy. Realistically, formal education generally cannot reach this level, at least in our society. In all open and pluralistic societies, such as our own, the maturity and personal integration required at this level are not attained until at least some years after the individual has completed his formal education. Time and experience must interact with affective and cognitive learnings before the individual can answer the crucial questions, "Who am I?" and "What do I stand for?" ³²⁰]

In addition, it will be recalled from p.39 that "Attitude, as does Value, spans Levels 2.2 through to Level 4.1", hence in an attitude assessment instrument based on the Krathwohl Taxonomy need not extend as far as level five.

The decision not to add a third response category to the 'true/not true' format - for example, 'uncertain' or 'neutral' - was made with regard to the hierarchical element of the Instrument. To explain, the item statements grouped into the four ascending levels of the Instrument's hierarchy were designed to elicit a correspondingly deeper level of commitment on the part of the respondents. An 'out' was already implicit in the lower level item statements, for example, words such as "I sometimes...", or 'I usually...' were utilized.

In the case of the higher level items, the statements were phrased more categorically. In other words, for the requirements of level four of the Instrument, one either 'always' follows a particular code of practice, or one doesn't. One either believes in some point strongly, or one doesn't. The 'uncertain' option at this point is not needed.

All item statements were analysed and modified where necessary according to the same criteria. The criteria adopted were as suggested by Allen Edwards in Techniques of Attitude

³²⁰ D. Krathwohl, B. Bloom, and B. Masia, (1964), op. cit., p. 165.

Scale Construction³²¹. In particular, at this early stage, many statements needed to be revised that:

- (1) were factual, or capable of being interpreted as being factual, or
- (2) needed shortening, or
- (3) were in the form of compound or complex sentences, or
- (4) used words that were unlikely to be understood by those who were to be given the completed Instrument, or
- (5) used double negatives.

The actual formulation and modification of item statements were tasks requiring close attention, particularly as the Instrument was intended for use with younger students (twelve to fourteen year olds). As indicated in one investigation, the way the individual items have been phrased may be crucial to the overall effectiveness of the Instrument. It was reported:

*The data from the present study provide strong evidence that the insertion of the word "not" has a profound influence on student responses. Two trends were indicated in the results. First, items that induced a more favourable response on the positive form induced a less favourable response on the negative form. In other words, respondents were less likely to indicate agreement by disagreeing with a negatively phrased item than to indicate agreement by agreeing with a positively phrased item. Second, items that induced an unfavourable response on the positive form were less likely to induce an unfavourable response on the negative form. ... A likely explanation of these findings is that elementary school children do not understand negation and, consequently, fail to convey their true attitude when confronted with a negatively phrased item.*³²²

³²¹ A. Edwards, Techniques of Attitude Scale Construction, (1957).

³²² J. Benson and D. Hocevar, "The Impact of Item Phrasing on the Validity of Attitude Scales for Elementary School Children", Journal of Educational Measurement, (1985), pp. 236-237.

The view of the present study with regard to the negative phrasing of statements is in accord with the above comments. The instrument being developed was intended for adolescent-age students, hence the need for simplicity. The occasional negative item was, however, to be included, so as to locate any respondents who may have ticked all the 'trues' or 'not trues'. These items were to be as simply stated as possible. [Refer to items 31, 32, 36, and 39, Appendix I]

The issue of semantics formed a major part of the development of this New Music Attitude Assessment Instrument (hereafter 'NMAAI'), with four separate drafts being examined before the final version was settled on [see p.134 for further details]. Many of the problems in the early stages were centred on the item statements of Level Four, a fact that is significant when one remembers that Lewy reported similarly: he found in his Music test that "...items pertaining to level 4 are more "popular" than those of level 3."³²³ In the present study the problem was isolated as being the lack of commitment required by many of the level four item statements: it was virtually too 'easy' to make a positive response. This will be further explained in the following discussion, which details the development of each item statement from its original form through to the final version. The level one item statements will be examined first, followed by those of level two, three then four. The number beside each item statement is the place in which it appears in the final Instrument [see Appendix I]. It will be noted that items drawn from past Instruments remain at their original level.

Statement Development:

Level one: Receiving (Attending)

The following ten items are at this level due to the fact that they are low in emotional content and are generally 'indications only' of interest.

³²³ A. Lewy, (1971), op. cit., p. 364.

(2) *I would like to improve my understanding of music.*

This statement was originally -

I have never had the wish to improve my understanding of music. (Lewy)

- which needed simplifying, and turning into the positive form, so as to be less confusing for the respondents.

(5) *I am usually aware of changes in loudness and softness in music.*

This statement was originally -

Do you notice changes in loudness and softness in listening to music?

(Noble)

- which required changing into the true/not true format. The word 'notice' was replaced with 'aware of'.

(11) *I would like to know more about differences between styles in music.* (Lewy)

This statement remained unchanged from the original Lewy Instrument.

(15) *I would like to know more about the music I listen to.*

This statement was originally -

I would like to know more about the history of music. (Lewy)

- which tended to be interpreted by adolescents as referring to the history of 'classical' music only. Thus an irrelevant issue was being 'tested', resulting in an inordinate number of negative responses (87% in the trialling of the adapted Lewy Instrument).

(17) *I would like to spend more time listening to music.*

This statement was originally -

I would like to devote some time to listening to good music. (Lewy)

- which contained an ambiguous qualitative component in the word 'good'. It was firstly amended to -

I would like to devote some more time to listening to music.

- then gained one further modification with the substitution of the word 'spend' for 'devote'.

(21) *I would like to know more about writing music down.*

This statement was originally -

I would like to know more about musical theory. (Lewy)

- but the resistance to the word 'theory' virtually guaranteed a negative response from all potential respondents (98% in the trialling of the adapted Lewy Instrument).

(26) *There are certainly special kinds of music I like to listen to.*

This statement was originally -

Are there any special kinds of music you like to listen to? (Noble)

- which merely required changing into the true/not true format.

(28) *I am often aware of my feelings when I listen to music.*

This statement was originally -

Are you aware of your feelings brought about by the music that you hear?

(Noble)

- which required firstly changing into the true/not true format:

I am often aware of my feelings brought about by the music that I hear.

The statement was then rewritten into a more concise version. (These changes were made for the first draft of the NMAAI.)

(29) *I would like to learn how to play a/another musical instrument, if I had the opportunity.*

This statement was originally -

Would you like to learn how to play a musical instrument, if you had the opportunity? (Noble)

- which required changing into the true/not true format, as well as the addition of the word 'another'.

(34) *In movies, on television, and on video, I am usually aware of the background music.*

This statement was originally -

In movies and on TV, are you consciously aware of the existence of the background music to the actors' words and actions? (Noble)

- which was firstly changed into the true/not true format. It was also considered necessary to add the choice of 'video', due to its increased availability and popularity. Hence (for the first draft of the NMAAI):

In movies, on TV, and on video, I am not usually aware of the background music to the actors' words and actions.

For the final draft of the Instrument, this statement had its 'positive' form reinstated.

Level two: Responding

The ten items at this level aim to uncover whether respondents are actually *seeking* musical involvement (i.e. some element of self-motivation). The level of commitment, however, is still not expected to be high, although the respondent is beginning to exhibit a positive attitude towards music in general.

(1) *Sometimes when I'm tired, it "refreshes" me to listen to some music.*

This statement was originally -

Sometimes when I'm tired, it "refreshes" me to listen to some short musical piece. (Lewy)

- however the phrase 'short musical piece' was considered unnecessary.

(4) *If I have been impressed by a musical piece, song etc., I usually look for other opportunities to hear it again.*

This statement was originally -

If I have been impressed by a musical piece, I usually look for additional opportunities to hear it again. (Lewy)

- however there was concern that the phrase 'musical piece' might be interpreted too narrowly by the majority of adolescents. Hence the addition of 'song etc'; there was also a change from 'additional' to 'other' in the interests of simplicity. (These changes were made for the first draft of the NMAAI.)

(7) *I go to concerts or other live musical performances at least once in a year.*

This statement was originally -

I do attend concerts at least once in a year."(Lewy)

- however it was felt that the term 'concerts' was too restricting and culturally biased on its own. Hence:

I do attend concerts or other live musical performances at least once in a year.

The words 'I do attend' was subsequently replaced with 'I go to', which is language more readily used by adolescents. (These changes were made for the first draft of the NMAAI.)

(22) *Music provides me with entertainment and relaxation.*

This statement was originally -

Music often provides me with entertainment and relaxation. (Lewy)

- however the word 'often' was considered unnecessary.

(31) *I have never written, or had the urge to write, a song/piece of music of my own.*

This statement was originally -

Have you ever had the urge to compose a melody of your own? (Noble)

- which was firstly amended to include the option of 'already composed'. The true/not true format was adopted next, as was a negative phrasing of the statement. It was considered important to include some 'negative' statements in the Instrument, so that respondents did not fall into a conditioned pattern. In the final version of the statement, the word 'never' appears early, so that respondents are prepared from the outset. Hence:

I have never composed, or had the urge to compose, a song/piece of music of my own.

The term 'compose', however, was found to have links with 'classical' music in the minds of adolescents, hence the further change in this statement for the final draft of the NMAAI.

(32) *At social gatherings, I don't usually join in the group singing.*

This statement was originally -

At social or church gatherings, do you usually join in the group singing?

(Noble)

- which was rewritten into the true/not true format, had the words 'or church' removed, and was changed to the negative form.

(36) *Although I will listen to music when it happens to be playing, I do not actively seek it out for myself.*

This statement was originally -

Although I often listen to music when I find it at hand, I do not actively seek it out for myself. (Hevner)

- but it required a simple change from 'when I find it at hand' to 'when it happens to be playing'.

(37) *Playing an instrument or singing is a good way for a person to spend leisure time.*

This statement was originally -

Playing an instrument or singing is a good way to spend leisure time.
(Original)

- but after initial trials it was discovered that some respondents thought it referred to themselves specifically. In the final draft of the NMAAI, the addition of the words 'for a person' ensured a general interpretation of the statement.

(39) *I am involved with music only when I have to be.*

This statement was originally -

I am involved with music only when I'm compelled to be."(Original)

- but in the interests of simplicity, the word 'compelled' was substituted with 'have'.

(40) *I enjoy associating with others in some musical activity.*

(Original)

This statement remained unchanged throughout the trials.

Level three: Valuing

At this level, items were selected that required a greater emotional involvement: e.g. respondents needed to ascribe some value to a musical phenomena, or indicate that they commit time freely to musical activity. It will be recalled that key words at this level were consistency, stability, and commitment.

(3) *Listening to music that I think is good is a part of my everyday life.*

This statement was originally -

Listening to good music is a part of my everyday life. (Lewy)

- but was amended so as to lessen the ambiguity of the phrase 'good music'.

(9) *I often read articles which deal with recently composed music.*

This statement remained unchanged from the original Lewy Instrument.

(13) *I like to discuss with my friends topics related to music.*

This statement was originally -

It is not unusual for me to talk about topics related to music. (Lewy)

- which was improved by removing the double negative.

(14) If I like a certain musical piece/song etc., I often recommend it to my friends.

This statement was originally -

If I like a certain musical piece, I often recommend it to my friends for the purpose of listening to it.”(Lewy)

- however the last seven words were deemed superfluous, and ‘musical piece’ was expanded to include ‘song etc’. (These changes were made for the first draft of the NMAAI.)

(16) I usually read newspaper reviews of new recordings.

This statement remained unchanged from the original Lewy Instrument.

(18) I attend concerts or other live musical performances quite regularly (on the average at least once in six weeks).

This statement was originally -

I do attend concerts quite regularly (on the average at least once in six weeks). (Lewy)

- which was expanded firstly with ‘or other live musical performances’:

I do attend concerts or other live musical performances quite regularly (on the average at least once in six weeks).

From the second draft of the NMAAI, the word ‘do’ was omitted.

(19) *When I listen to a record, I like to read the notes that accompany it.*

This statement was originally -

I like to listen to records which contain explanations or discussion of the music played. (Lewy)

- which was shortened and made more direct. Nevertheless, in one of the final trials, a respondent wrote: "I can't read music" - in other words, he understood 'notes' to mean musical notes, not prose.

(23) *To offer a complete education, a school needs to include music.*

This statement originally underwent a number of revisions, as follows:

A curriculum which does not include music does not offer a complete education. (Original)

A school which does not include music does not offer a complete education. (Second draft of the NMAAI.)

The language finally selected is simple, the statement is direct, and the concepts involved are broad, without being personal to the respondent.

(27) *It is usual for me to plan in advance any musical programme I shall listen to on the radio or television.*

This statement was originally -

It is very unusual for me to plan in advance which musical program I shall listen to on the radio in the coming week. (Lewy)

- which needed modifying on two counts: (1) removing the use of the double negative, and (2) adding the choice of the television.

(38) *Music brings people together and fosters school and community spirit.* (Original)

This statement remained unchanged throughout the trials.

Level four: Organization

The items of this level were chosen for their potential to challenge more deeply the respondent's involvement with music. Evidence of a 'value system' by which the respondent appraises musical concepts and pursuits is sought.

(6) *I sometimes consider how a particular type of music reveals something about the people who created it.*

This statement was originally -

Have you ever tried to explain or discuss the following topic: How a particular type of music reveals something about the people who created it?
(Lewy)

- and required firstly changing into a single sentence:

I have often considered how a particular type of music reveals something about the people who created it.

One further modification was necessary, in order to make the statement more immediate to the respondent. (This change was made for the final draft of the NMAAI.)

(8) *When listening to music, I usually try to make a judgement about the quality of the performance.*

This statement was originally -

When listening to music, I usually try to make judgements about the quality of the performance. (Lewy)

- but here, the idea of 'commitment' was missing. The answer was to change the words 'make judgements' to 'make a judgement', and the statement took on a much more significant focus.

(10) *I discuss the quality of performance with others who have heard the same programme/record.*

This statement was originally -

I like to exchange views on the quality of performance with others who have heard the same program. (Lewy)

- which was extended to include 'record':

I like to exchange views on the quality of performance with others who have heard the same program/record.

As this is a level four statement, however, the words 'I like to' were in practice rather too 'watered down'. The final version is more decisive, and requires more of a commitment from the respondent who replies 'true'.

(12) I try to assess the elements that make a musical composition/song etc. a classic.

This statement was originally -

Have you ever tried to explain or discuss the following topic: What makes a musical composition immortal? (Lewy)

- which was, in the first 'modified Instrument' trial, altered for the benefit of the young respondents to:

Have you ever tried to explain or discuss the following topic: What makes a musical composition/song etc. a classic?

The word 'classic' was found to be readily understood by adolescents, as it is frequently used with regard to popular music. However, following the first trial, the form of this question had to be changed into the true/not true format, hence:

I sometimes wonder what makes a musical composition/song etc. a classic.

Yet again though, the impact of this statement as belonging to level four of the taxonomy was lost, and it was discovered that more commitment was required by the respondent when the words 'I sometimes wonder what makes' were replaced by 'I try to assess the elements'.

(20) In forming an opinion on performance, I am critical of the technique of the performers.

This statement was originally -

In forming an opinion on performance, I consider the technique of the performers. (Lewy)

- but it gained more significance when the words 'I consider' were replaced by 'I am critical of'. (This change was made for the first draft of the NMAAI.)

(24) I like to compare and evaluate performances of the same music by different musicians.

This statement was originally -

It is not unusual for me to compare performance of the same music by different musicians. (Lewy)

- which required the removal of the double negative:

It is usual for me to compare performances of the same music by different musicians.

Once again greater commitment was required by the respondent, hence the addition of the words 'and evaluate'.

(25) I am interested in comparing musical styles from different historical periods.

This statement was originally -

Have you ever tried to explain or discuss the following topic: What are the differences between the musical styles of various historical periods? (Lewy)

- which was changed into the true/not true format for the first draft of the NMAAI:

Discussing the differences between the musical styles of various historical periods doesn't interest me particularly.

This was subsequently shortened and simplified to:

I am not interested in comparing musical styles from different historical periods.

On further consideration, the word 'not' was removed from this statement.

(30) I am interested in the way that music can generate or express feelings.

This statement was originally -

Have you ever tried to explain or discuss the following topic: How music can create or express feelings? (Lewy)

- which was firstly changed into the true/not true format:

I am interested in the way that music can create or express feelings.

The substitution of the word 'generate' for the word 'create' was one further modification.

(33) In my opinion, everyone should take some music at school.

This statement was originally -

In schools we should concentrate on subjects that relate directly to jobs, rather than on Arts and recreational subjects. (Original)

- which was revised in the interests of simplicity. These are both level four statements, but the final version (introduced from draft two of the NMAAI) is more effective by being directed to the respondent personally.

(35) Professional performing groups, either pop or classic, are important enough for me to spend my own money on tickets to attend their concerts.

This statement was originally -

Professional performing groups, both pop and classic, are integral to the maintenance and development of our musical culture. Because of the financial difficulties almost all of them are presently having in this country, they should be directly subsidized by the federal government, such as groups are in Europe. (Noble)

- which was immediately rewritten into a more concise form for the first draft of the NMAAI:

Professional performing groups, both pop and classic, are integral to our country's musical culture: therefore they should be directly funded by the government.

The problem with this statement, however, is that it did not appeal directly to the adolescent: for example, most young people are not particularly concerned about Government funding. It was too easy, therefore, to give a positive response without being totally committed to the concepts being tested. The final version of this statement is extremely personal, and although some respondents complained of a lack of finance, the statement was always marked 'true' by students who had, overall, a highly positive relationship with music.

3. TESTING GROUPS

Respondents involved in the present research were primarily New Zealand secondary school students, predominantly from the ages of twelve to sixteen years old. These comprised the main testing groups, the groups isolated for reliability and validity studies, and the smaller groups examined for specific reasons (to be discussed). The remaining testing groups were from the University of Canterbury (Christchurch, New Zealand), or were employees of the Wellington Harbour Board.

With regard to the origins of the main respondent groups, four different secondary schools were involved, these being located in three New Zealand cities: Auckland (1 school), Wellington (1 school), and Christchurch (2 schools). The Auckland school was used in the pilot study, and is co-educational. The first of the Christchurch schools (hereafter 'Chch 1') was the other school to be used in the pilot study, but it was involved also in the testing of the NMAAI. 'Chch 1' was another co-educational school. The second of the Christchurch schools (hereafter 'Chch 2') may be described as belonging to a higher socio-economic bracket than does 'Chch 1', and it is a public school for girls only. The Wellington school is a private Catholic college for boys only. The respondents were obtained through convenience sampling.

For comparative reasons, and in order to examine various hypotheses, the respondents were identified as falling into distinct categories. These categories will now be described.

Pilot study:

As has already been discussed, the group involved with the testing of the adapted Lewy Affective Assessment Instrument consisted of 75 third and fourth form general music students (12 to 15 year olds). New to this discussion, however, but undertaken during the pilot study, was the administration of the modified Lewy Instrument to fourteen university students, who were studying Music Education in the second year of their Music degrees. This was the precursor to the validity studies, which will be outlined in due course (beginning on p.135). The outcome of this administration of the Lewy Instrument will be discussed with other results in Chapter VI.

The NMAAI:

Main Respondent Groups.

The NMAAI passed through three drafts before being presented to the first group of respondents. These were nineteen third form general music students from the first of the two Christchurch schools ('Chch 1') involved with the research. Following this presentation, some further modifications were made to the Instrument, resulting in the fourth and final draft.

Details of the main respondent groups used in the administration of this final draft of the NMAAI are as follows:

<u>Chch 1</u>	
3 rd form	81
4 th form	35
TOTAL	116

<u>Chch 2</u>	
3 rd form	29
4 th form	27
TOTAL	56

<u>Wellington 1</u>	
3 rd form	75
4 th form	63
TOTAL	138

<u>Totals</u>	
Total 3 rd form	185
Total 4 th form	125
Total tested	310

Reliability and Validity.

Reliability:

Reliability studies were undertaken as a separate exercise using a group of 3rd form students previously untested with the NMAAI.

First, the test-retest procedure was selected, with 92 respondents completing the NMAAI a second time following a ten-day interval. Test-retest reliability coefficients were obtained using the Pearson Product-Moment Correlation method applied to the two total scores (out of forty) of each student.

Second, the Kuder-Richardson formula for estimating reliability was applied to the data obtained from the second administration of the NMAAI (N=92). This result was used to calculate the standard error of measurement (SEM), so as to be able to establish confidence intervals for individual scores.

Validity:

In order to ascertain the validity of the NMAAI, a number of additional studies and comparisons were undertaken. These consisted of:

1. the administering of the NMAAI to a group of 3rd and 4th form students observably demonstrating a positive attitude toward musical activity,
2. the administering of the NMAAI to a group of university music students, and
3. the comparing of teacher, and student self ratings with NMAAI performance.

1. For the first validity study, the NMAAI was administered to a group of 'Chch 1' 3rd and 4th form students who were participating in an extra-curricula production of a musical stage show. As these students had volunteered to devote their free time to staging such a large-scale musical event, it was reasoned that their attitude towards musical activity would be high. Ideally then, the performance of those students on an Instrument designed to assess that attitude level should be correspondingly high. The total number of students tested was thirty.

In order to analyse the resultant data in terms of statistical significance, the null hypothesis - *there is no difference between the NMAAI mean score of the Musical Production group and the NMAAI mean score of the Main Respondent group* - was assumed. So as to support the use of the Musical Production group in a validity study, it was desired that the null hypothesis be rejected.

2. In furthering the validity examination, both the adapted Lewy Instrument and the NMAAI were administered to university students [fourteen each Instrument] who were studying Music Education in the second year of their Music degrees. [see p.133 under 'Pilot Study'] These students were selected as being of value to a validity study because they had demonstrated behaviorally - by their decision to study Music Education at a tertiary level - the existence of a positive affective relationship with music. Obviously several factors may have influenced their subject choice at University, but one ought to be able to safely assume that a positive affective relationship with music was one of them.

A second null hypothesis - *there is no difference between the NMAAI mean score of the University Music group and the NMAAI mean score of the Main Respondent group* - was stated in this instance. This was also expected to be rejected.

The reasoning behind this last exercise was that the performance of the university students on the music attitude assessment Instruments ought to be notably high; however there was one specific variable that needed to be isolated before any conclusions from these results

could be drawn. That variable was 'age of respondent'; in other words, it had to be established whether university music students performed significantly well on the attitude assessment Instruments simply by virtue of their being 'older' than the adolescent respondents. To offset the University group then, the NMAAI was administered to a group of 24 adults 'with no specific musical associations'. These adults were all employees of the Wellington Harbour Board (with full time occupations such as office workers, and marine and helicopter pilots), and had not demonstrated through their post secondary school activities any particular 'musical association'. All in fact verbally endorsed this last aspect of themselves, despite the fact that some had learned to play musical instruments, and/or were otherwise involved with music in the past.

3. The final exercise undertaken to complete the validity studies involved the comparing of teacher ratings of students, and students' own self-ratings, with NMAAI performance. The respondent group used for this was the test-retest group (N=92), each student being asked to rate their own attitude to music (on a scale of 1-10) on the two occasions that they completed the NMAAI. Independently, the music teacher also submitted ratings, representing his understanding of each student's attitude to music. All ratings were compared with the 'total score' performance of the students on the NMAAI. The Pearson Product Moment Correlation Coefficient was calculated to measure the relationships in each instance.

Smaller Testing Groups.

As has been mentioned, smaller testing groups fulfilled specific roles in the study.

To begin with, it will be recalled that one of the criteria of the NMAAI was to address the issue of cultural unfairness, and to this end, separate data were collated from the responses of students who identified themselves as being of predominantly ethnic origin. In New Zealand, this group (extracted from the Main Respondent group, and the 'test-retest'

group) consisted mainly of Maori and Pacific Island students. The details of these respondents are as follows:

<u>Chch 1</u>	
Maori	34
Others	3
Total	37

<u>Wellington 1</u>	
Maori	9
Others	8
Total	17

<u>Totals</u>	
Total Maori	43
Total Others	11
Total tested	54

The reason for isolating the above group was to compare statistically the performance of the ‘ethnic’ students with the performance of the Main Respondent group as a whole. The desired outcome of this exercise was to show that this particular sub-grouping of students offers patterns of data which parallel those of the ‘mainstream’. This would indicate that the NMAAI is ‘culturally fair’.

The null hypothesis, that *there is no difference between the NMAAI mean score of the Ethnic group and the NMAAI mean score of the Main Respondent group* was stated in this instance. So as to support the notion that the NMAAI is ‘culturally fair’, it was desired that the null hypothesis would *not* be rejected in a test of statistical significance.

Further to an examination of the aims of the NMAAI, another small respondent group was isolated: this time in order to consider one aspect of the ‘cognitive versus affective assessment’ issue. This group comprised both 3rd and 4th form students from ‘Chch 1’ who had previously been classified as being ‘slow learners’ with regard to their cognitive/intellectual abilities. Once again the aim was to compare statistically their performance with that of the Main Respondent group. The numbers involved are as follows:

<u>Chch 1</u>	
3 rd form	8
4 th form	9
Total ‘slow’ learners	17

The null hypothesis, that *there is no difference between the NMAAI mean score of the Slow Learner group and the NMAAI mean score of the Main Respondent group* was stated in this instance. It was desired that this would not be rejected also, following analysis of the data.

A summary of the various respondent groups is as follows:

Figure 2: Summary of NMAAI respondent groups

Group		School etc	Number
General music (main group)	3 rd form	ChCh 1	81
	4 th form		35
	3 rd form	ChCh 2	29
	4 th form		27
	3 rd form	Wgtn 1	75
	4 th form		63
General music (test-retest group)	3 rd form	ChCh 1	92
Musical Production	3 rd and 4 th form	ChCh 1	30
University Music Students		University of Canterbury	14
Adults/ no musical association		Wellington Harbour Board	24
Ethnic	3 rd and 4 th form	ChCh 1	37
	3 rd and 4 th form	Wgtn 1	17
Slow Learner	3 rd and 4 th form	ChCh 1	17

4. RESPONDENT INFORMATION SHEET

Each of the approximately 440 respondents involved with the final draft of the NMAAI was asked to complete a 'Respondent Information Sheet' [see Appendix K] so as to facilitate the collection of certain data for statistical purposes. This exercise primarily enabled the researcher to identify the 'ethnic' group, as respondents were given the opportunity to state their ethnic background. Questions I and III were basically a 'foil' for the ethnic background question, although some interesting observations were made following an examination of the 'music preference' answers.

Predictably, perhaps, the most commonly preferred types of music for the respondents were 'pop' and 'rock' genres; however, following an examination of the completed questionnaires it became apparent that adolescents do discriminate decisively about their music. [This is supported by Hirsch, who suggests that "the stratified teenage audience (usually viewed by adults as an undifferentiated horde) is an aggregate of individuals who form distinct popular music subaudiences."³²⁴]

³²⁴ P. Hirsch, "Sociological Approaches to the Pop Music Phenomenon", American Behavioral Scientist 14 (1970-71), p. 379.

A survey of 200 respondent information sheets yielded the following outcomes:

MUSIC TYPES	NO. OF RESPONDENTS CHOOSING THIS OPTION*
Rock	119
Pop	108
Heavy Metal	54
Classical (this was often 'qualified' e.g. 'light' or 'some')	36
Rap	30
Country and Western	19
Reggae	11
Rock 'n Roll	9
Blues	7
Jazz	6
60s Music	6
Modern Music	3
Country	3
Hard/Heavy Rock	3
Folk	2
'Weird' Music	2
Soul	2
Funk	2
Hippy	2
All sorts	2
Any except country & western	2

* NOTE: Respondents were not restricted in their choice of music preferences.

Other descriptions suggested by single respondents were opera, Christian, old-fashioned, house, mellow, 'any except classical', death metal, 'olden day songs', Maori, 'up to date' music, and 'anything my parents don't listen to'.

CHAPTER VI - RESULTS OF THE NMAAI TRIALS

1. BACKGROUND

It will be recalled that an affective assessment instrument, developed in the USA, and thought to be of potential value to the present study, was modified and updated, and readministered to New Zealand secondary school students. The results obtained were favourable: from a group of 75 respondents, a total of 73% "...showed a tendency toward a hierarchy pattern". A second aspect to be tested was found to be positive also, that "consensus can be obtained among expert judges as to the proper taxonomic level of given behavior."

This paved the way for the development and testing of a new Attitude Assessment Instrument (the NMAAI), which was given an initial trial when in its third draft form. [Refer to Chapter V, where details of the respective drafts are given.] For each completed Instrument the number of 'positive' responses to item statements at each level was totalled, giving rise to a 'four score' result. The particular pattern formed by each of these was then categorised as they were in the Pilot Study, as follows:

- A = 'Scores' gained at each taxonomic level follow a distinct hierarchy pattern (For example, '8-6-4-3').
- B = Slight deviation from hierarchy pattern, i.e. 'scores' from three of the four taxonomic levels represented follow a hierarchy pattern (For example, '7-8-4-3', or '8-6-3-4').
- C = There is more than one disruption to the hierarchy pattern (for example, '6-8-3-4').

In addition, two further categories were included, D and E; the first, D, (which was to become important as far as the scoring of the NMAAI was concerned) was initiated by

Robert Noble in his notion of a respondent 'meeting a level'. With reference to the administration of the original version of his Instrument, he stated:

*...if the respondent's opinions agreed with those the jury perceived as "appropriate responses" to four out of five inventory items measuring a level, the respondent was adjudged to have "met" that level. ... Inventory items were not grouped together by hierarchical level; rather, they were rearranged to avoid patterns in either questions or responses.*³²⁵

In order to 'meet' any of the four hierarchical levels of the NMAAI, respondents were required to give 'positive' responses to at least seven out of the ten item statements measuring each level. This 'cut off point of seven', whilst being (initially) arbitrary, was decided upon after close examination of the patterns of data as they emerged during the trialling of the NMAAI. Essentially, a 'cut-off point of eight' was found to be too stringent, as data collected under this regime failed to show strong discriminatory patterns. Ten item statements were included at each taxonomic level of the instrument so as to allow for variations in respondents' attitudes; a 'cut off point of seven' for each level offered flexibility in this respect, whilst still providing a 'standard' for 'meeting a level'.

Each completed Instrument then, as well as generating a 'four score' (scores being out-of-ten) result, was also given a 'meeting the levels' profile, which consisted of four 'met' or 'had not met' outcomes. In other words, if a respondent had, for example, gained a 'four score' result of 8-7-5-3, they would have 'met' levels one and two, but not levels three and four, as would have a respondent with a 'four score' result of 7-8-5-3.

The 'cut-off point of seven' category may also be examined for the existence of hierarchical patterns; for example, the respondents from above who 'meet' levels one and two but not

³²⁵ R. Noble, "A Multivariate Analysis of Factors in the Backgrounds of Wyoming Adults Related to their Attitudinal Levels Concerning Music", A Cross-Section of Research in Music Education, ed: S. Barnes, (1982), p. 233.

levels three and four have contributed hierarchical patterns, whereas a respondent with the 'four score' result of 7-5-8-3 has not. Interestingly, a 'four score' result such as 3-5-4-6 would be interpreted as a hierarchy pattern under the 'cut off point of seven' category (as the 'meeting the levels' profile would consist of four 'have not met' outcomes), but not under the original hierarchy category (category 'A', from above) calculated directly from the 'four score' result.

The second additional category, E, included in the analysis of the NMAAI data arose out of the concern that, of all the levels represented in the Instrument, Level Four was emerging as the one most likely to give problems statistically. (It will be recalled from Chapter V [see p.116] that "...Lewy reported similarly: he found in his Music test that "...items pertaining to level 4 are more "popular" than those of level 3."³²⁶ In the present study the problem was isolated as being the lack of commitment required by many of the level four item statements: it was virtually too 'easy' to make a positive response.") In other words, of all the 'four score' results that did not follow a hierarchical pattern, a significant number 'offended' by collecting a higher score at Level Four than were gained at at least two of the other levels. (For example, a 'four score' result of 7-4-3-6 would come under this description.)

It was decided to monitor this trend throughout the trialling of the NMAAI, and attempt to redress the problem by further modifying the Level Four item statements. This last category, then, identifies the percentage of 'four score' results that collect "...a higher score at Level Four than were gained at at least two of the other levels" (from above).

All together, the five categories under which the data from the trialling of the NMAAI were initially analysed were as follows³²⁷:

³²⁶ A. Lewy, (1971), op. cit., p.364.

³²⁷ The notion of hierarchy is addressed in further statistical detail in the section headed 'Scalogram Analysis, p.165.

- A = 'Scores' gained at each taxonomic level follow a distinct hierarchy pattern (For example, '8-6-4-3').
- B = Slight deviation from hierarchy pattern, i.e. 'scores' from three of the four taxonomic levels represented follow a hierarchy pattern (For example, '7-8-4-3', or '8-6-3-4').
- C = There is more than one disruption to the hierarchy pattern (for example, '6-8-3-4').
- D = The 'meeting the levels' profile (taken from the 'cut-off point of seven' calculation) follows a distinct hierarchy pattern (for example, a 'four score' result of 8-7-4-3 would translate into a 'meeting the levels' profile of 'has met Level One - has met Level Two - has not met Level Three - has not met Level Four' - which constitutes a hierarchy pattern).
- E = The 'four score' result displays a higher score at Level Four than was gained at at least two of the other levels.

In short, the desired outcomes according to the above classification scheme would be as follows:

- Category A }
 Category B } a high percentage of the data [85% & above]
- Category C : a very low percentage of the data [15%& below]
- Category D : a high percentage of the data
- Category E : a very low percentage of the data

The results of the trialling of the NMAAI may now be reported; these will in turn be interpreted in Chapter VII, "Conclusion: Interpretations and Implications".

2. RESULTS

Main Respondent Groups.

The third draft of the NMAAI was administered to nineteen third form students from the ‘Chch 1’ school, the results being as follows:

Table 3: NMAAI results - ‘ChCh1’ 3rd form

GROUP	No.	A	B	C	D	E
Chch 1 (3 rd form)	19	16%	68%	16%	58%	21%

From Table 3 it can be seen that a total of eighty-four percent (A plus B) “showed a tendency toward a hierarchy pattern”, comparing well with the seventy-three percent result obtained during the pilot testing of the adapted Lewy Instrument. Of a more unsatisfactory nature were the results of categories A and D, which one would want to see much higher; and E, which needed to be reduced significantly. Hence these results - bearing in mind the small numbers of the sample group - were used to direct the further modification of the item statements of the NMAAI. The details of this are outlined in Chapter V.

From these modifications emerged the fourth draft of the NMAAI, which was trialled with the following results:

Table 4: NMAAI results - ‘Chch 1’ 3rd and 4th form

GROUP	No.	A	B	C	D	E
Chch 1 (3 rd form)	81	33%	48%	19%	73%	12%
Chch 1 (4 th form)	35	17%	72%	11%	69%	11%
Together	116	28%	55%	17%	72%	12%

Table 5: NMAAI results - ‘Chch 2’ 3rd and 4th form

GROUP	No.	A	B	C	D	E
Chch 2 (3 rd form)	29	41%	48%	10%	83%	10%
Chch 2 (4 th form)	27	19%	59%	22%	70%	19%
Together:	56	30%	54%	16%	77%	14%

Table 6: NMAAI results - ‘Wgtn 1’ 3rd and 4th form

GROUP	No.	A	B	C	D	E
Wgtn 1 (3 rd form)	75	36%	51%	13%	73%	7%
Wgtn 1 (4 th form)	63	41%	51%	8%	76%	13%
Together:	138	38%	51%	11%	75%	9%

Table 7: NMAAI results - All 3rd and 4th form

GROUP	No.	A	B	C	D	E
Total 3 rd form:	185	36%	49%	15%	75%	10%
Total 4 th form:	125	30%	58%	12%	73%	14%
Total tested:	310	33%	53%	14%	74%	11%

Of interest was the similarity of the overall results obtained from the three trial schools - despite the geographical and socioeconomic differences. Table 8 illustrates this trend:

Table 8: NMAAI results - A comparison of all schools

GROUP	No.	A	B	C	D	E
Chch 1 3 rd & 4 th Form:	116	28%	55%	17%	72%	12%
Chch 2 3 rd & 4 th Form:	56	30%	54%	16%	77%	14%
Wgtn 1 3 rd & 4 th Form:	138	38%	51%	11%	75%	9%
Total 3 rd Form:	185	36%	49%	15%	75%	10%
Total 4 th Form:	125	30%	58%	12%	73%	14%
Total Tested:	310	33%	53%	14%	74%	11%

These results, particularly with regard to the development of the NMAAI from the third to the fourth draft, were found to be very promising.

Table 9 presents the data from a different perspective, allowing for a closer examination of the effects of the ‘third to fourth draft’ modifications on the Categories A, D and E:

Table 9: The development of the NMAAI from 3rd to 4th draft

CATEGORY	% Range	Average%	%Attained
	Draft Four	Draft Four	Draft Three
A	17-41	33	16
B	48-72	53	68
C	8-22	14	16
D	69-83	74	58
E	7-19	11	21

As a result of the trialling of the third draft of the NMAAI, it will be recalled that the percentage of responses falling into Category A (in other words, the percentage of ‘four score’ results following a distinct hierarchy pattern) was unsatisfactorily low (16 percent). Table 9 indicates the success of the fourth draft modifications: of the six groups tested, the

range of percentages falling into Category A was 17-41, with the mean for the whole group (of 310 respondents) being 33 percent.

Similar improvements are noted in Category D (the percentage of 'meeting the levels' profiles which follow a hierarchy pattern, following the 'cut off point of seven' calculation). The draft three result was 58 percent, as opposed to 74 percent obtained from the trialling of the fourth draft of the NMAAI. This was considered to be a positive increase.

Category E was the third category under special observation - the aim being to reduce the percentage of 'four score' results that collected "...a higher score at Level Four than were gained at at least two of the other levels". Following the close analysis and subsequent modification of Level Four item statements, the situation appeared to have been redressed: the fourth draft of the NMAAI yielded a final percentage in Category E of 11 percent (with a range of 7 - 19), decidedly more satisfactory than the 21 percent obtained from the trialling of Draft Three.

Reliability and Validity.

Reliability:

As the group used for the reliability studies was previously untested with the NMAAI, full details of the administration results were collected so as to compare the group as a whole with the Main Respondent group. The results - recalling that the Instrument was administered twice - were as follows (remembering from p.146 that these will be discussed further in Chapter VII):

Table 10: NMAAI results - A comparison of the Test-retest reliability group with the Main Respondent group

GROUP	No.	A	B	C	D	E
Test - retest group, 1 st administration	92	30%	59%	11%	71%	8%
Test - retest group, 2 nd administration	92	33%	57%	10%	70%	5%
Main Respondent Group	310	33%	53%	14%	74%	11%

An additional statistic was examined in order to further the comparison between the test-retest and the Main Respondent group. This involved calculating the ‘score out of forty’ - that is, the total number of positive responses from each returned NMAAI was recorded (in this instance, the hierarchical nature of the NMAAI was discounted). These scores were totalled and then averaged for each group:

Table 11: NMAAI results - A Comparison of Mean Scores
- Test-Retest and Main Respondent Groups

GROUP	No.	Mean	Standard Deviation
Test - Retest group, 1 st administration	92	24.7	6.35
Test -Retest group, 2 nd administration	92	25.0	7.22
Main Respondent Group:	310	25.3	7.00

It was on the ‘scores out of forty’ that the coefficient of reliability was calculated, using the Pearson Product-Moment Correlation method, that being:

$$r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2] [N\sum Y^2 - (\sum Y)^2]}}$$

The result obtained with the NMAAI Test - Retest group was $r = .90$.

The 'scores out of forty' were also used in the calculations involving the Kuder-Richardson formula for estimating internal consistency reliability, that formula being:

$$r = \frac{(n)}{(n-1)} \times \frac{(SD^2 - \sum pq)}{(SD^2)} \quad \text{where } n = \text{no. of items in the test}$$

p = proportion 'passing' an item (in this case it was 'responding positively' to an item)
q = 1 - p

With the present data, the above became:

$$r = \frac{(40)}{(39)} \times \frac{(7.22^2 - 7.38)}{(7.22^2)}$$

$$r = 1.02 \times \frac{(52.13 - 7.38)}{(52.13)}$$

$$r = 1.02 (0.86)$$

$$r = 0.88$$

This reliability estimate was then used to "calculate the standard error of measurement (SEM), so as to be able to establish confidence intervals for individual scores" (p.135). Those calculations are as follows:

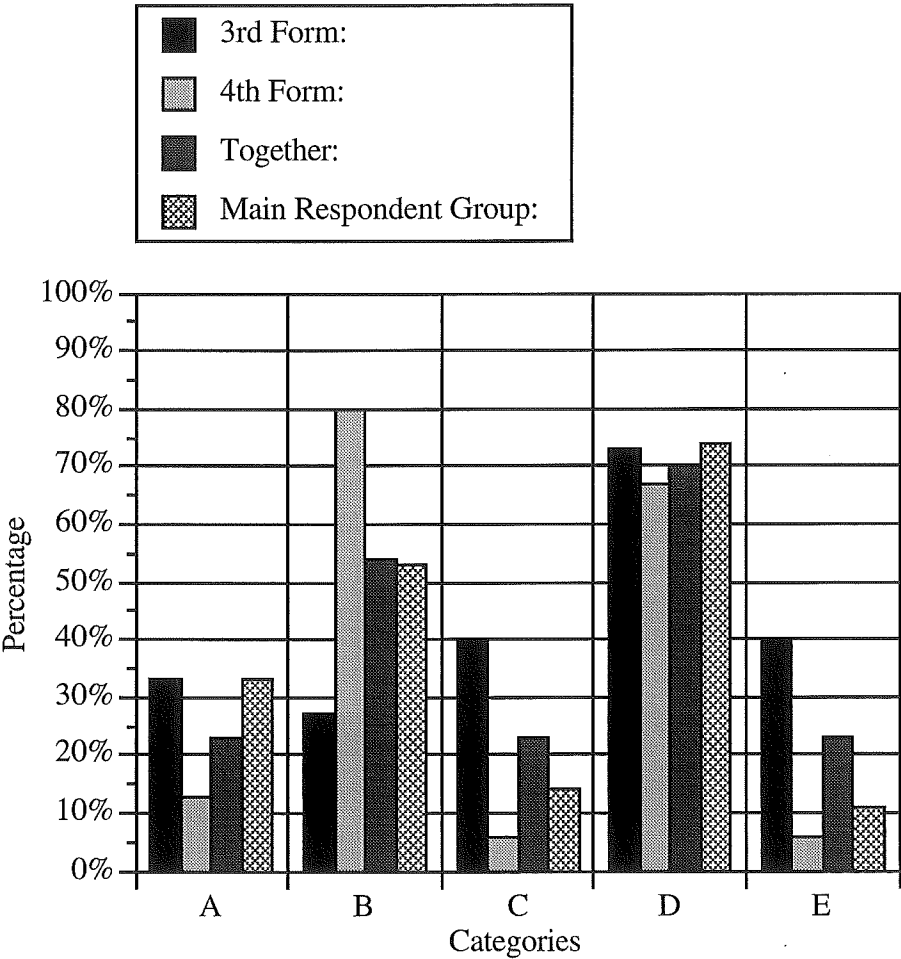
$$\begin{aligned} \text{SEM} &= SD \times \sqrt{[1-R]} \\ &= 7.22 \times \sqrt{[1- .88]} \\ &= 7.22 \times .35 \\ &= 2.53 \end{aligned}$$

These results will also be discussed further in Chapter VII.

Validity:

1. The first respondent group to be administered the NMAAI for validity purposes was the ‘Chch 1’ musical production students. The data collected from this group are as follows:

Figure 3: NMAAI results - A Comparison of Musical Production and Main Respondent Groups



The average (mean) ‘score out of forty’ for the Musical Production group, as it compares with that of the Main Respondent group, appears in Table 12:

Table 12: NMAAI results - A Comparison of Mean Scores
- Musical Production and Main Respondent Groups

GROUP	No.	Mean	Standard Deviation	SE _{mean}	95% Confidence Interval
Musical Production Group, 3 rd Form:	15	29.4			
Musical Production Group, 4 th Form:	15	30.5			
Musical Production Group, together:	30	29.97	5.3	0.97	28.07 to 31.87
Main Respondent Group:	310	25.3	7.0	0.40	24.52 to 26.08
MR Grp. without Production Grp.*:	280	24.78	6.44	0.39	24.02 to 25.54

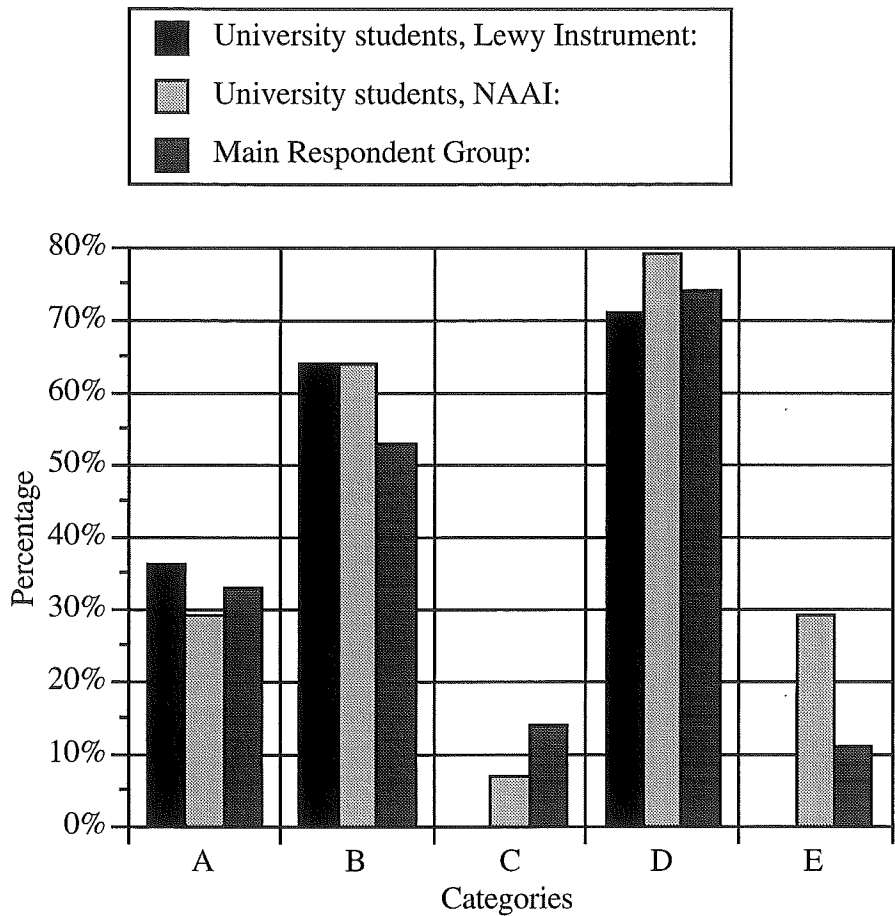
* The Musical Production Group respondents were also in the Main Respondent Group.

In order to ascertain the significance of the results, the *t* test (using the separate variance formula) was applied. The *t* value obtained (with respect to the Musical Production Group, *n*=30, and the Main Respondent Group with the Musical Production Group extracted, *n*=280) was 4.76. The *average* value taken from the *t* distribution table at the 0.05 level is 2.0; thus it may be concluded that the obtained *t* = 4.76 is significant³²⁸. Indeed, even at the 0.0005 level, (where the average *t* value = 3.5) the value of 4.76 is larger: hence the result is still significant.

2. Further to the examination of the validity of the NMAAI, fourteen university music students completed the instrument (as outlined on page 136, Chapter V), and these results were compared with 1) the results of a similar group who had completed the modified Lewy Instrument some weeks before, and 2) the results of the Main Respondent group (from above). The data are reported in Figure 4:

³²⁸ The procedure for averaging tabled *t* values is as recommended by L. Cohen and M. Holliday, Statistics For Social Scientists, [1982], pp. 232-233.

Figure 4: NMAAI results - A Comparison of University Music and Main Respondent Groups



The average (mean) ‘score out of forty’ for the University Music Students group, as it compares with that of the Main Respondent group, appears in Table 13:

Table 13: NMAAI results - A Comparison of Mean Scores
- University Music and Main Respondent Groups

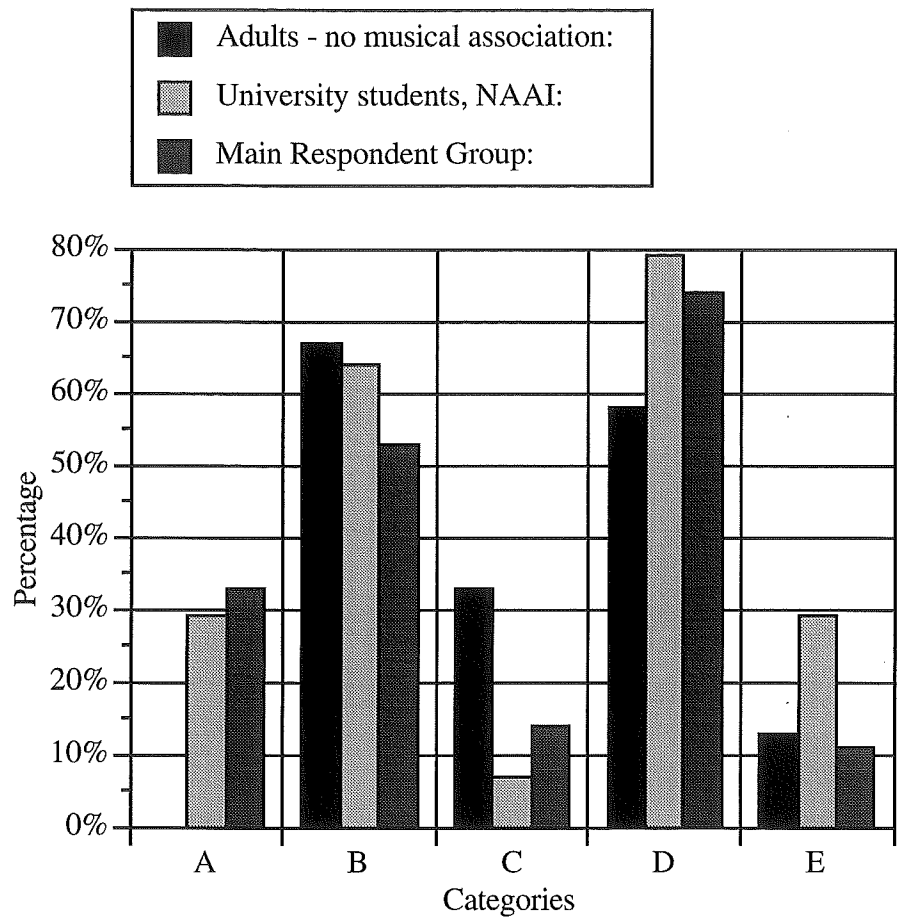
GROUP	No.	Mean	Standard Deviation	SE mean	95% Confidence Interval
University Students (NMAAI):	14	34.5	4.7	1.27	32 to 37
Main Respondent Group:	310	25.3	7.0	0.40	24.52 to 26.08

On applying the t test (using the separate variance formula), the t value obtained in this instance was 6.97. The *average* value taken from the t distribution table at the 0.05 level is 2.06; thus it may be concluded that the obtained $t = 6.97$ is significant. Once again, even at the 0.0005 level, (where the average t value = 3.75) the value of 6.97 is larger: hence the result is still significant.

Other interesting statistics resulting from a comparison of the University Music Students group with the Main Respondent group concern the 'meeting the levels' profiles. Of the fourteen respondents in the University Students group, 71% met the criteria of Levels One through Four (in other words, had made at least seven positive responses in each of the four taxonomic levels of the NMAAI), whereas in a group of one hundred third form students from the Main Respondent group, only 23% matched that performance.

It will be recalled that a second adult group was involved in the testing of the NMAAI, so as to offset the 'age' factor possibly influencing the results of the University Students group. This group consisted of twenty-four adults 'with no specific musical associations' (see p.137), and their results were as follows:

Figure 5: NMAAI results - A Comparison of Adult and Main Respondent Groups



The average (mean) ‘score out of forty’ for the ‘adults with no particular musical association’ group, as it compares with that of the University Students group and the Main Respondent group, appears in Table 14:

Table 14: NMAAI results - A Comparison of Mean Scores
- Adult and Main Respondent Groups

GROUP	No.	Mean	Standard Deviation	SE mean	95% Confidence Interval
Adults - no musical association:	24	25.6	5.8	1.18	23.3 to 27.9
University Students (NMAAI):	14	34.5	4.7	1.27	32 to 37
Main Respondent Group:	310	25.3	7.0	0.40	24.52 to 26.08

The mean difference between the two groups in the table (Adults and University Students Groups) is statistically significant ($p<0.0001$).

3. The third validity study, involving the comparison of student and teacher ‘attitude to music’ ratings, focussed on a 1-10 scale [refer p. 137]. The actual ‘total score’ performances of the students on the NMAAI were converted so as to correspond also to this scale (for example, a respondent who scored $\frac{40}{40}$ would receive a ‘10’; one who scored $\frac{20}{40}$ would be a ‘5’, and so on). As the respondents used for this exercise belonged to the test-retest group, two sets of self- and NMAAI ratings were gathered in each instance. The teacher rated the students once only. The initial results are as follows:

- 1. 85% of the respondents rated themselves more highly than their NMAAI ratings; another 11% rated themselves similarly. Only 4% gave themselves a lower rating than they received from their NMAAI performance.
- 2. The teacher rated the students lower than their NMAAI rating 62% of the time; they received a teacher rating similar to their NMAAI rating 13% of the time. Thus 25% of respondents received a higher teacher than NMAAI rating.
- 3. At the second completion of the NMAAI, 87% of the respondents gave themselves a rating that was within one point of their first self-rating.

Hence the emerging data tended to form a pattern consisting of NMAAI ratings falling between a higher self-rating on the one side, and a lower teacher-rating on the other (67% of cases). For example:

Self-Rating 1	Self-Rating 2	NMAAI Rating 1	NMAAI Rating 2	Teacher-Rating
7	7	5	4	3
8	8	6.5	7	5
9	9	6	6	5

The correlation coefficients, calculated to measure the nature of the relationships concerned, were found to be as follows:

- i) Teacher-ratings v NMAAI ratings (1st run) - r=.11
- ii) Teacher-ratings v Self-ratings (1st run) - r=.30
- iii) Self-ratings (1st run) v Self-ratings (2nd run) - r=.81
- iv) Self-ratings (1st run) v NMAAI ratings (1st run) - r=.30
- v) NMAAI ratings (1st run) v NMAAI ratings (2nd run) - r=.89

These results will be interpreted in Chapter VII.

Smaller Testing Groups.

In addition to the above, the NMAAI was administered to various smaller testing groups (detailed in Chapter V). The first of these, it will be recalled, was the group comprising "...students who identified themselves as being of predominantly ethnic origin." The data collected from this group are as follows:

Table 15: NMAAI results - Chch 1 Ethnic Group

GROUP	No.	A	B	C	D	E
Chch 1 Maori:	34	29%	59%	12%	74%	3%
Chch 1 Others:	3	33%	33%	33%	100%	nil
Together:	37	30%	57%	13%	74%	3%

Table 16: NMAAI results - Wgtn Ethnic Group

GROUP	No.	A	B	C	D	E
Wgtn 1 Maori:	9	33%	56%	11%	67%	nil
Wgtn 1 Others:	8	50%	25%	25%	75%	25%
Together:	17	41%	41%	18%	71%	12%

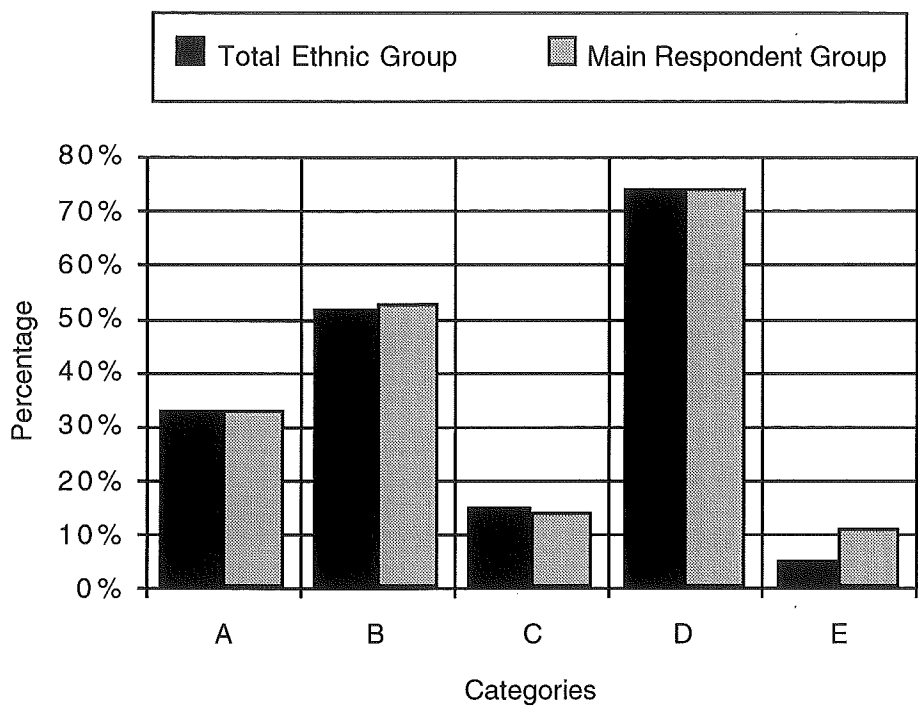
One may regroup the above data into the following categories:

Table 17: NMAAI results - Total Ethnic Group

GROUP	No.	A	B	C	D	E
Total Maori:	43	30%	58%	12%	72%	2%
Total Others:	11	45%	27%	27%	82%	18%
Total Tested:	54	33%	52%	15%	74%	5%

It is interesting at this point to compare directly the final results of the Ethnic group with those of the Main Respondent group; the implications will be discussed in Chapter VII:

Figure 6: NMAAI results - A Comparison of Ethnic and Main Respondent Groups



The average (mean) ‘score out of forty’ for the Ethnic group, as it compares with that of the Main Respondent group, appears in Table 18:

Table 18: NMAAI results - A Comparison of Mean Scores
- Ethnic and Main Respondent Groups

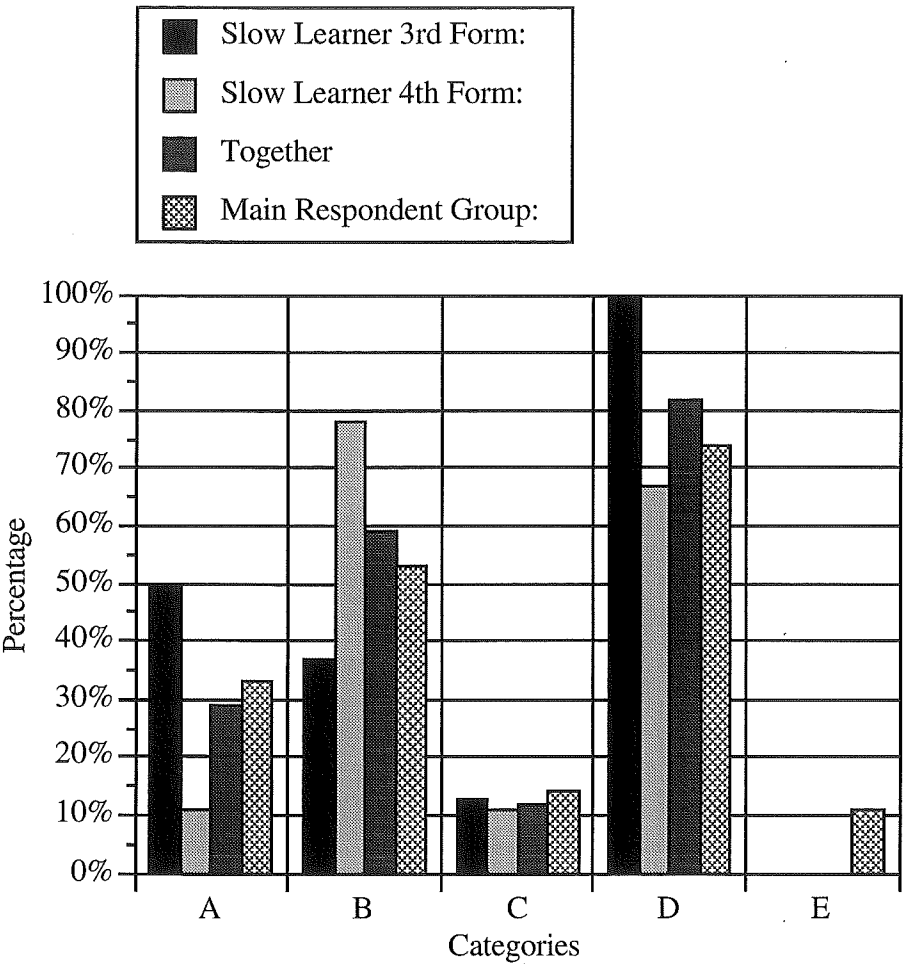
GROUP	No.	Mean	Standard Deviation	SE _{mean}	95% Confidence Interval
Maori:	43	25.8			
Others:	11	26.2			
Total Ethnic:	54	25.8	5.7	0.78	24.27 to 27.33
Main Respondent Group:	310	25.3	7.0	0.40	24.52 to 26.08
MR Grp. without Ethnic Grp.*:	279	25.25	6.54	0.39	24.49 to 26.01

* 31/54 of the Ethnic Group respondents were also in the Main Respondent Group.

In order to ascertain the significance of the results, the t test (using the separate variance formula) was applied. The t value obtained (with respect to the Total Ethnic group, $n=54$, and the Main Respondent group with the Ethnic group extracted, $n=279$) was .63. The *average* value taken from the t distribution table at the 0.05 level is 1.98; thus it may be concluded that the obtained $t = .63$ is not significant.

The second small respondent group to be isolated for comparison consisted of seventeen third and fourth form students from the 'Chch 1' school "...who had previously been classified as being 'slow learners' with regard to their cognitive/intellectual abilities." The results from this group are shown in Figure 7:

Figure 7: NMAAI results - A Comparison of Slow Learner and Main Respondent Groups



The average (mean) ‘score out of forty’ for the Slow Learner group, as it compares with that of the Main Respondent group, appears in Table 19:

Table 19: NMAAI results - A Comparison of Mean Scores
- Slow Learner and Main Respondent Groups

GROUP	No.	Mean	Standard Deviation	SE _{mean}	95% Confidence Interval
Slow Learner 3 rd Form:	8	25.8			
Slow Learner 4 th Form:	9	23.6			
Total Slow Learner:	17	24.7	7.5	1.8	21.17 to 28.23
Main Respondent Group:	310	25.3	7.0	0.40	24.52 to 26.08
MR Grp. without Slow L. Grp.*:	293	25.23	6.8	0.40	24.45 to 26.01

* The Slow Learner Group respondents were also in the Main Respondent Group.

On applying the *t* test (using the separate variance formula), the *t* value obtained in this instance was .28. The *average* value taken from the *t* distribution table at the 0.05 level is 2.04; thus it may be concluded that the obtained *t* =.28 is not significant.

A summary of the average (mean) ‘score out of forty’ for the various respondent groups is as follows:

Table 20: NMAAI results - A Summary of Mean Scores

- All Respondent Groups

GROUP	No.	Mean	Standard Deviation
Main Respondent:	310	25.3	7.0
Test-Retest Group (1st NMAAI)	92	24.7	6.35
Test-Retest Group (2nd NMAAI)	92	25.0	7.22
Musical Production Group, 3 rd Form:	15	29.4	
Musical Production Group, 4 th Form:	15	30.5	
Musical Production Group, together:	30	29.97	5.3
University Students (NMAAI):	14	34.5	4.7
Adults - no musical association:	24	25.6	5.8
Maori:	43	25.8	
Others:	11	26.2	
Total Ethnic:	54	25.8	5.7
Slow Learner 3 rd Form:	8	25.8	
Slow Learner 4 th Form:	9	23.6	
Total Slow Learner:	17	24.7	7.5

3. SCALOGRAM ANALYSIS

The NMAAI was developed according to the hierarchical principles of the Affective Taxonomy of Krathwohl *et al.*. Much of the work that was undertaken in the trialling of the Instrument focussed on the patterns and relationships which resulted from combining affective behaviours in music with those hierarchical principles. Ultimately, the most pressing question to ask is how successful has the NMAAI been in incorporating a hierarchical structure - is it a valuable tool for music education?

In answering this question one needs make a specific analysis of the performance of the NMAAI Levels as they relate to the criteria of a single continuum or cumulative scale. To explain, Allen Edwards defines a cumulative scale as being one from which "...an

individual with a higher rank (or score) than another individual on the same set of statements must also rank just as high or higher on every statement in the set as the other individual. In the case of attitude statements, we might say that this means that a person with a more favourable attitude score than another person must also be just as favourable or more favourable in his response to every statement in the set than the other person.”³²⁹

For the purposes of the present discussion, this definition may be adapted to the effect that, for example, any individual meeting the criteria of Level Three must also have met the criteria required by Levels One and Two; by the same token, ‘Level Four’ respondents must have exhibited and surpassed the performances forwarded by those attaining Levels One, Two, or Three. We need to establish from an analysis of the NMAAI trials results whether the above definition holds, bearing in mind the statement of Edwards that:

*... perfect scales exist only as ideal models and in practice it is necessary to determine the extent to which the data or observed patterns of response fit the model of a perfect scale.*³³⁰

Hierarchy.

There are different methods of testing for hierarchy. Two approaches have been used in this study, namely, Goodenough’s method of scalogram analysis (as described by Edwards, above), and a variation on the Binomial Test for determining the significance of a proportion. Each of these will now be discussed.

Goodenough:

Goodenough’s method of scalogram analysis (as described by Edwards) was selected in order to determine the coefficient of reproducibility of the NMAAI. With a cumulative

³²⁹ A. Edwards, (1957), op. cit., p. 172.

³³⁰ Ibid, p. 184.

scale, the coefficient of reproducibility represents "...the degree of accuracy with which we can reproduce the responses to statements from total scores alone."³³¹ In the case of the present study, the coefficient of reproducibility would indicate the extent to which we could assume, for example, that a Level Four respondent (in other words, a respondent who has made positive responses to at least seven out of the ten item statements measuring that level) had also 'met' the criteria required by Levels One, Two and Three.³³² The data for this analysis are found in Category D, which was defined on page 146 as follows:

D = The 'meeting the levels' profile (taken from the 'cut-off point of seven' calculation) follows a distinct hierarchy pattern (for example, a 'four score' result of 8-7-4-3 would translate into a 'meeting the levels' profile of 'has met Level One - has met Level Two - has not met Level Three - has not met Level Four' - which constitutes a hierarchy pattern).

Each 'meeting the levels' profile is expressed as a series of '1's and '0's - a '1' denotes that the respondent has met a particular level, a '0' that the respondent has not. To illustrate, the example cited above of a 'four score' result of 8-7-4-3 would be expressed as 1-1-0-0, meaning that Levels One and Two had been met, but not Levels Three and Four. Similarly, a 'four score' result of 6-5-7-2 would be expressed as 0-0-1-0, meaning that only Level Three had been met.

³³¹ Ibid.

³³² There are links between this information and the Category D percentages, however the former results from a more rigorous procedure for analysing the data. The category D percentages are concerned solely with the results which actually do follow a hierarchy pattern; whereas the coefficient of reproducibility is calculated from the deviations from the hierarchy pattern of all the results.

The coefficient of reproducibility is calculated on the proportion of errors which occur when the 'meeting the levels' profiles do not follow an ideal pattern. With regard to the hierarchical nature of the NMAAI, the ideal patterns are as follows:

0-0-0-0

1-0-0-0

1-1-0-0

1-1-1-0

1-1-1-1

According to the procedure outlined by Edwards³³³, the coefficient of reproducibility was calculated first using the various divisions of the Main Respondent group, and then using the entire sum of respondents involved in the trialling of the NMAAI. The results of this appear in Table 21 :

³³³ A. Edwards, (1957), op. cit., pp. 184-188.

Table 21: Coefficients of Reproducibility

Group	No.	Coeff. of R.
'Chch 1' 3 rd Form	81	.87
'Chch 1' 4 th Form	35	.84

'Chch 2' 3 rd Form	29	.91
'Chch 2' 4 th Form	27	.85

'Wgtn 1' 3 rd Form	75	.86
'Wgtn 1' 4 th Form	63	.88

Total 3 rd Form	185	.87
Total 4 th Form	125	.86

Total 3 rd & 4 th Form	310	.87
--	-----	-----

University Students	14	.89
---------------------	----	-----

Total Respondents	324	.87
-------------------	-----	-----

This Table will be discussed in Chapter VII.

Binomial Test for Significance of a Proportion:

One way of viewing the data collected in the present study is to examine each person's response to determine the *extent* to which it displays hierarchy. For example, if a respondent's four-score result was 8-6-5-3, it can be said to have displayed perfect hierarchy in that the subscores are progressively smaller as the levels ascend (i.e. $8 > 6 > 5 > 3$). This respondent has no conflicts or discontinuities in her/his results.

By contrast a second respondent who scores, for example, 8-5-6-3 can be said to have *one* discontinuity, namely that the third level score is greater than the second (i.e. $8 > 5 < 6 > 3$).

A further respondent who scores 8-3-6-5 has *two* discontinuities, namely that the second level score is lower than both the third and fourth.

By extending this method of analysis, and ignoring for the moment ‘ties’ in the scores (e.g. 8-5-5-3), the following discontinuities with distributions are possible:

Table 22:

DISCONTINUITIES	FREQUENCY BY CHANCE
0	1
1	3
2	5
3	6
4	5
5	3
6	1
TOTAL POSSIBILITIES:	24

In order to justify the claim that the data should satisfy hierarchy, one would expect that a significantly greater proportion of respondents would have discontinuities of 2 or less (i.e. showing *some* level of hierarchy), than discontinuities of 3 (*no* pattern of hierarchy) or 4 or more (some level of *reverse* hierarchy).

One complicating factor is the occurrence of ‘ties’: for example a four-score result of 8-8-5-3 incurs a discontinuity of $1\frac{1}{2}$, because of the tie between the first and second level scores. Similarly a four-score result of 6-5-5-5 incurs a discontinuity of $1\frac{1}{2}$.

Allowing for ties, the following distribution of discontinuities was obtained for the present data:

Table 23:

DISCONTINUITIES	N	%
0-1½	74	24
1-1½	110	35
2-2½	76	25
3-3½	36	12
4-6	14	4

From these results it can be seen that 84% of respondents’ scores displayed some level of hierarchy (a discontinuity of 2½ or better). However it could be argued that discontinuities of 2 and 2½ are somewhat marginal in terms of showing that hierarchy is present. For this reason, discontinuities of 1½ or better have been chosen as indicating clear evidence of hierarchy; - as can be seen in Table 23, 59% of cases reached this stronger level.

Statistically, this percentage is significant using a modification of the Binomial Test for Proportion. If one were to define hierarchy as YES (discontinuities = 0 - 1½) and NO (discontinuities = 2+), and it is argued that there is at least an equal probability of obtaining a YES or NO (in fact, statistically, there is a greater chance of scoring NO: approx 0.79³³⁴), one would expect 155 of the main respondent group to score YES and 155 to score NO.

Using the Binomial Test ³³⁵.

$$Z = \frac{x - np}{\sqrt{npq}}$$

where n=sample size; p=proportion expected for YES; q=1-p

³³⁴ Probability value supplied by Dr Cedric Hall, University Teaching and Development Centre, Victoria University of Wellington.

³³⁵ C. Leach, Introduction to Statistics, (1979), p. 135.

- it was calculated:

$$Z = \frac{184 - (0.5 \times 310)}{\sqrt{[310 \times 0.5 \times 0.5]}}$$

$$Z = \frac{184 - 155}{\sqrt{77.5}}$$

$$Z = \frac{29}{8.803}$$

$$Z = 3.294$$

This result, $Z = 3.294$, is associated with a p value of 0.0005, based on the Normal Curve Distribution. It should be stressed that this is a conservative test for establishing the significance of the presence of hierarchy in the data; it will be recalled that the present calculations are based on the assumption that, by chance, the probability of gaining a discontinuity score of $1\frac{1}{2}$ or better was 0.50. In fact, the probability was 0.21.

The intention in this instance was to be deliberately conservative in order to claim with conviction that hierarchy exists.

CHAPTER VII - CONCLUSION: INTERPRETATIONS AND IMPLICATIONS

1. SUMMARY OF THE INTRODUCTORY SECTION (Chapters I to IV)

The aim of the introductory section of the present study was to identify and examine the issue of affective domain assessment. In particular, attention was focussed on locating general problem areas, predominantly from a music education viewpoint. Nine specific 'difficulties and problems associated with assessment in the affective domain' were presented and discussed.

The affective taxonomy of Krathwohl, Bloom and Masia was introduced, and discussion centred on the respective elements of the cognitive and affective domains. An investigation was made into the difficulties which arise when one attempts to make divisions between the two, with the emerging question being one of 'means and ends'. In other words, "the problem occurs ... when the cognitive elements assume an all-consuming importance, and the affective 'ends' are relegated to mere byproducts of the educational process." (pp.21-22)

Three specific questions, relating to music education, were considered, the intention being to investigate the need for undertaking further research. The first two were as follows:

- (1) Why focus at all on affective elements in assessment?; and
- (2) Why test specifically for attitude levels? (Can one not simply assume from a student's level of academic/cognitive achievement a related attitudinal level?)

Essentially, it was found that affective elements such as attitude and interest, being integral to the music experience (and therefore worthy in their own right of independent evaluation), did not necessarily follow the same path of development as their cognitive

counterparts. This conclusion was drawn from examining the literature and research of prominent educationalists (see pp.26-30).

The actual direction adopted by the present study was prompted by the third question, that being:

- (3) Why is there a need for structuring an approach to the assessment of affective elements?

The answer in this instance is that too often the evaluation of affective domain elements is hampered by their seeming intangibility. However, as Colwell states:

The fact that cognition is a legitimate part of the affective response to music helps simplify the problems of measuring affective response. Evaluation of attitudes and values, though intricate, is possible because many aspects can be verbalized, put into specific statements, and clearly pointed out by teacher or student. ³³⁶

Hence the advantage of looking to the affective taxonomy of Krathwohl, Bloom and Masia, as this model was intended as a means of formally “ordering and relating the different kinds of affective behavior.”³³⁷

The ‘Purpose of the Study’ chapter resumed the examination of the problems arising from affective domain assessment. A more in-depth investigation was undertaken of the affective Taxonomy of Krathwohl *et al*, specifically outlining the different hierarchical levels of the model. In particular, the position of ‘attitude’ on the Taxonomy continuum was highlighted.

³³⁶ R. Colwell, (1970), op. cit., p.129.

³³⁷ D. Krathwohl, B. Bloom, and B. Masia, (1964), op. cit., p. 24.

The deficiencies of existing affective domain assessment - as outlined in Chapter I - were considered, and suggestions were made in each instance as to possible solutions to be incorporated into an attitude assessment instrument. The success of these solutions, in the light of the NMAAI trials, will be re-examined at a later point in this chapter. [see pp.187-191]

Chapter III, 'The Affective Component of New Zealand Secondary School General Music Education', sought to "identify the affective component of modern-day New Zealand secondary school music education ... thereby describing the circumstances which frame the current quest to develop an attitude assessment instrument." (p.52) An historical approach was adopted, briefly tracing the development of secondary school music education in this country. The latest music education syllabus was described in some detail, with an investigation being made into the status that was being granted to affective domain elements. The investigation concluded that "affective domain objectives are now accorded official recognition in the new music education syllabus for schools", but that "the procedures for their assessment are inadequately addressed." (p.70)

The 'Review of the Research' section investigated studies which focussed on affective assessment and music in general. With attitude tests and other affective assessment instruments in music, two categories emerged, those being 1) tonal (containing those tests which employ musical stimuli), and 2) verbal (comprising those which "demand either retrospective or prophetic valuation because they do not use actual music"). Apropos these categoric divisions, the discussion was firstly divided into two parts: 1) Studies and research directly related to 'The Affective Domain and Assessment in Music' (as defined by the present study) [focussing primarily on verbal test literature], and 2) Studies and research (into affective assessment in music) which are not directly related to the present topic [in other words, referring to tonal test literature.] The latter was included "so as to further delimit the topic in hand." (p.72)

The third section of the 'Review of the Research' comprised an examination of studies which involve the affective Taxonomy in some way; firstly those with general application to educational concerns, and secondly those principally relating to music education. Interestingly, all research concerning the affective taxonomy had a favourable outcome. This supported, and was consistent with, the findings of the present study.

All of the above succeeded in setting the stage for the development of the 'New Music Attitude Assessment Instrument': the NMAAI.

2. DISCUSSION AND INTERPRETATION OF THE NMAAI TRIALS (Chapters V and VI)

In Chapter V, the development and trialling of the NMAAI was outlined, with the results being reported in Chapter VI. A summary of those results will now be presented, along with interpretations where appropriate.

Pilot study:

As part of the present research, prior to the development of the NMAAI, a past affective assessment instrument was modified and updated, and then readministered in its new form to groups of secondary school students. The outcomes of this pilot study were reported and discussed at the beginning of Chapter V (see p.104-112), and a brief review may be found in Chapter VI (p.143).

The NMAAI:

The item statements for the new instrument, the NMAAI, were drawn from a variety of sources (see p.112), and "...were chosen particularly for their appropriateness and relevance to today's students." (p.112) These statements were given considerable attention

throughout the trials - with particular attention being paid to respondent feedback - and were modified wherever necessary. The final form of the NMAAI (see Appendix I) was regarded as satisfactory in terms of the level of commitment engendered by each statement and its allocated taxonomic level. In addition, comprehension difficulties appeared to have been reduced to a minimum. [Following each administration of the instrument, the supervisor questioned the students informally with regard to this matter.]

The resultant data from the trialling of the NMAAI were initially organised into five categories [see Chapter VI, p.146]. Of major interest were categories C, D, and E; as in order for the NMAAI to be successful, C and E needed to show very low percentages [the goal being below 15%], and D high ones [the goal being above 85%].

Essentially, Category D contained the most important statistic, that being the percentage of 'meeting the levels profiles' (taken from the 'cut off point of seven' calculations) which follow a distinct hierarchy pattern. The reason for the importance of this statistic is that one is looking, in the results, for the emergence of the hierarchy principle on which the NMAAI is based. To clarify:

- (1) The theoretical model (the Taxonomy of Krathwohl *et al*) is structured hierarchically.
- (2) The NMAAI was developed according to the hierarchical principles of that model.
- (3) The results obtained from administering the NMAAI should point to the underlying hierarchy, particularly with regard to the four 'levels' of the Taxonomy.

Hence if '3' does not occur, then either the Taxonomy (with regard to its suitability for assessment in Music Education), or the NMAAI - or both - are in need of further investigation (Chapter V, p.104 saw the initiation of this issue). In either case, support would not be found for the theories expounded in Chapter I of the present study.

In the words of Lewy, then, one would wish to discover that activities "... at a higher taxonomic level will occur less frequently than those at a lower level" (see p.105, Chapter V). On achieving this objective, one may attribute the NMAAI with some success in reflecting the hierarchical feature of the Taxonomy. It will be noticed that the five categories for organising data focus on the 'levels' structure of the NMAAI, as opposed to the total scores gained by respondents.

Main Respondent Groups.

The development of the NMAAI from the third to the fourth drafts showed marked improvements in the percentages obtained for all categories. The importance of these was discussed on p.116. Of further interest was the similarity of the results obtained from the three trial schools - despite geographical and socioeconomic differences.

It was stated earlier that "...in order for the NMAAI to be successful, C and E needed to show very low percentages, and D high ones." (p.177) From Table 9 (p.149) it can be seen that the objective appears to have been achieved as far as Categories C and E are concerned. Category C, that being the percentage of 'Scores' gained at each taxonomic level not indicating the existence of a hierarchy pattern, yielded a result of 14% from a total of 310 students trialled. Category E, containing the percentage of 'four score' results which display a higher score at Level Four than was gained at at least two of the other levels, totalled only 11%. Both Categories, it may be concluded, yielded markedly low percentages [the goal being below 15%] from the trial group of 310 students.

Category D, the percentage of 'meeting the levels' profiles (taken from the 'cut-off point of seven' calculation) which follow a distinct hierarchy pattern, culminated with a result of 74% (with the individual groups ranging from 69% - 83%). While this statistic is a distinct improvement from the 58% achieved in the trialling of the Third draft of the NMAAI, it is not sufficient to register as a 'high' percentage [the goal being above 85%] with regard to

the success of the NMAAI. Certainly, though, there is evidence that the hierarchy principle of the Taxonomy of Krathwohl *et al* has emerged in the NMAAI results. Further evidence may be found in the discussion beginning p. 186.

Categories A and B, those being (A), 'scores' gained at each taxonomic level follow a distinct hierarchy pattern and (B), 'scores' from three of the four taxonomic levels represented follow a hierarchy pattern, remained consistent throughout the NMAAI fourth draft trials, capturing between them a total of 86% [this being considered a 'high' result, i.e. above 85%]. The total for these categories from the third draft trials was also 86%, however with the fourth draft a different distribution resulted in a heavier weighting for Category A.

Reliability and Validity.

Reliability:

An examination of the patterns of data received for the test-retest group (p.151) showed that on both administrations of the NMAAI the results were similar to those received for the Main Respondent group. This was important so as to be able to continue with the reliability coefficient calculations. If the two groups had been too dissimilar, generalisations would not have been possible with regard to the test-retest reliability coefficient.

As was reported, the coefficient of reliability calculated using the Pearson Product-Moment Correlation was .90 which was considered to be very favourable. In order to interpret this result, the following recommendation by Lehman may be considered:

A reliability of .85 is generally considered satisfactory, although it depends upon the nature of the test, the variability within the group, and the purposes for which the test is used. If a test is used to distinguish between individuals, it is desirable to have a relative coefficient of at least .90, but for

some group measurements reliabilities as low as .60 or even .50 may be useful. ³³⁸

From this evidence then, one could suggest that the level of reliability of the NMAAI is such that it will consistently discriminate amongst New Zealand secondary school students with respect to their test scores.

The estimate of reliability obtained by applying the Kuder-Richardson formula, whilst lower than that obtained in the test-retest exercise (.88 v .90), was still considered to be favourable. It suggests that the items in the test generally contribute to a single underlying trait.

The estimate of .88 was used in the calculation of a standard error of measurement of the NMAAI. As reported on p.152, the SEM was found to be 2.53: this means that, for example, if a respondent scores $30/40$, their 'true' score is:

30 ± 2.53 (27.47 to 32.53) - at the 68% confidence level

or

30 ± 4.96 (25.04 to 34.96) - at the 95% confidence level

Validity:

The results of the validity studies, despite the small numbers being tested, allow for some interesting comparisons.

1. With respect to the group, 'Musical Production Students, Chch 1 School', percentages in three of the five Categories paralleled those of the Main Respondent group, with Categories C and E being slightly high (see Figure 3, p.153).

The notable result occurred in the 'score out of forty' calculation, with the mean for the Musical Production group being $29.97/40$: a substantial increase from the $25.3/40$ mean of the Main Respondent group. The difference in the means was found to be significant even at the 0.0005 level, hence the null hypothesis, that *there is no difference between the NMAAI*

³³⁸ P. Lehman, (1968), op.cit., p.13.

mean score of the Musical Production group and the NMAAI mean score of the Main Respondent group was able to be rejected.

This shows that the 'Musical Production' students, on average, marked positively more item statements on the NMAAI than was average for their age group, according to the results of the NMAAI trials.

2. The second trial concerning the validity of the NMAAI involved the 'University Music Students' group. As the students had demonstrated a high affective relationship with music by undertaking Music Education studies at tertiary level, the aim was for this to be reflected in the results. Indeed, this aim was to be achieved:

Firstly, the percentages gained in the five categories successfully mapped the patterns of data of the Main Respondent group (see Figure 4, p.155).

Secondly, the average (mean) 'score out of forty' for the 'University Music Students' group (34.5) was substantially higher [as was to be expected] than that gained by the Main Respondent group (25.3). The difference in the means was found to be significant even at the 0.0005 level, hence the null hypothesis, that *there is no difference between the NMAAI mean score of the University Music group and the NMAAI mean score of the Main Respondent group* was able to be rejected.

A third area of comparison involved the 'meeting the levels' profiles for the two groups, as reported on p.156. The high percentage of University music students meeting the criteria of Levels One through Four ("in other words, had made at least seven positive responses in each of the four taxonomic levels of the NMAAI") far exceeded the percentage of Main Respondent group students.

Hence one may suggest that the NMAAI, by its performance in the above validity trial, has the potential to reflect the level of students' attitudes to music.

In order to "offset the 'age factor' possibly influencing the results of the University students group", twenty-four adults 'with no specific musical associations' were invited to complete the NMAAI. The patterns of data emerging from this group were interesting to compare with those of both the University music students group, and the Main Respondent group (see Figure 5, p.157). The most striking feature was the seemingly random nature of the results received:

Category C, that being 'Scores' gained at each taxonomic level do not indicate the existence of a hierarchy pattern, yielded the undesirably high percentage of 33; whilst Category D, the 'meeting the levels' profile (taken from the 'cut-off point of seven' calculation) follows a distinct hierarchy pattern, was well under the other statistics received at 58%.

The unpredictability of the results from the Adult group indicates a lack of compatibility between these respondents and the NMAAI. This is perhaps not surprising due to the fact that the Instrument was developed solely for use with secondary school-age students - yet it is interesting that the Instrument does appear to differentiate between 'adults with no specific musical associations' and adults who have demonstrated a high level of commitment to music. In other words, the NMAAI appears to offer greater reliability for respondents who are of secondary school age, or who are musically committed.

Further to a comparison of the Adult group with other groups tested, the average (mean) 'score out of forty' in the NMAAI was 25.6: marginally up on the Main Respondent group mean of 25.3, yet considerably lower than the University music students' mean of 34.5 (see Table 14, p.157). This result offers support for the use of the University music students in a validity context.

3. The third validity study, involving teacher- and self-ratings with the test-retest group, yielded some interesting disparities. It will be recalled that the initial results appeared to

follow a pattern consisting of “NMAAI ratings falling between a higher self-rating on the one side, and a lower teacher rating on the other (67% of cases)” (p.158).

In the 11 instances where the teacher assigned a rating for a student which differed *markedly* from this pattern, for example -

Table 24:

Self-Rating 1	Self-Rating 2	NMAAI Rating 1	NMAAI Rating 2	Teacher Rating
7	6	2	1	8
3	3	2	1.5	6
9	10	7.5	8.5	3

- there was found to be a personality or behavioural element involved (either the student was cognitively very ‘bright’ and hence appeared to attract a higher teacher rating than was warranted, or the student was perhaps disruptive in class, indicating a negative attitude to music as far as the teacher could ascertain. These explanations were submitted by an independent teacher).

It is interesting to note a comment by Lehman with regard to teacher ratings:

Studies have shown teachers’ evaluations to be dependent upon a number of factors that limit their effectiveness as criteria for validity tests. ³³⁹

Such factors may include the personality and behavioural elements mentioned above.

With respect to the reported product moment correlation coefficients (p.159), one may interpret the results according to a Cohen and Holliday ‘descriptive interpretation of *r*’ table³⁴⁰:

³³⁹ P. Lehman, (1968), op.cit., p.16.
³⁴⁰ L. Cohen and M. Holliday, Statistics For Social Scientists, [1982], p. 93.

1. there is a very low correlation between teacher-ratings and the ratings assigned to students from their NMAAI scores.
2. there is a low correlation between both students' self-ratings and teacher-ratings, and students' self-ratings and the ratings assigned to them from their NMAAI scores.
3. there is a high correlation between students' self-ratings on two separate occasions.
4. there is a high correlation between the ratings assigned to students from their NMAAI scores on two separate occasions.

It may be that subjectivity is the major factor which undermines the usefulness of teacher- and self- ratings in a validity context. This notwithstanding, one would have to conclude that the information gathered in this part of the present study does not enhance the validity of the NMAAI. Indeed, because of the subjective nature of this process and the very small number of teachers involved in this instance, further work would be needed in order to produce a valid coefficient for individual assessments.

Smaller Testing Groups.

The first of the smaller respondent groups to be tested consisted of the "students who identified themselves as being of predominantly ethnic origin". The objective was to "compare statistically the performance of the 'ethnic' students with the performance of the Main Respondent group as a whole. The desired outcome of this exercise was to show that this particular sub-grouping of students offers patterns of data which parallel those of the 'mainstream'. This would indicate that the NMAAI was "culturally fair." (Chapter II, p.45)

The results, as reported in Chapter VI pp.159-162, were very promising, with the final combined percentages for the 54 respondents of the Ethnic group matching very closely those of the Main Respondent group (see Figure 6, p.161). The greatest difference - of only six percent - occurred in Category E, with the Ethnic group showing the more desirable result. The average (mean) 'score out of forty' calculation for the Ethnic group

(which was not taking into account the hierarchical principle of the NMAAI) also compared favourably with the Main Respondent group. The results were $25.8/40$ for the former, and $25.3/40$ for the latter. The slight difference in the means was found to be not significant at the 0.05 level, hence the null hypothesis, that *there is no difference between the NMAAI mean score of the Ethnic group and the NMAAI mean score of the Main Respondent group* could not be rejected.

There were large variations, however, when one compared the two sections of the Ethnic group: the 'Maori' and the 'Others'. Differences of up to 31% were recorded, and it is not until all the results are combined in each category that the pattern comes into line with that of the Main Respondent group. The small testing numbers of the 'Others' group may be a contributing factor; nonetheless, further investigation would be needed to clarify this matter.

As far as the overall statistics are concerned, there is certainly an indication that the NMAAI may be culturally fair. There were minimal differences between the two groups with regard to 'patterns of data', this being, it will be recalled, the desired outcome of the exercise. Due to the small numbers being tested, however, one may make no further inferences other than that the NMAAI has the potential to stand up to further scrutiny in this area.

The results of the second small respondent group, the 'slow learner' third and fourth form students, were also encouraging with regard to a comparison of the patterns of data.

Figure 7 (p.163) shows a similarity of percentages in each category despite the very small numbers in the Slow Learner group as opposed to the Main Respondent group. The fact that the 'score out of forty' mean was lower for the Slow Learner group than for the Main Respondent group, was found to be not significant at the 0.05 level. The null hypothesis, that *there is no difference between the NMAAI mean score of the Slow Learner group and the NMAAI mean score of the Main Respondent group* could therefore not be rejected.

Hierarchy.

In seeking to answer the question “how successful has the NMAAI been in incorporating a hierarchical structure” an investigation was undertaken into the “performance of the NMAAI Levels as they relate to the criteria of a single continuum or cumulative scale”. [refer pp.165-172] By calculating the ‘coefficient of reproducibility’ for the NMAAI levels, one is able to estimate the likelihood that any ‘four-score’ result will follow a progressive hierarchical pattern. Ideally, one would want the coefficient to be .85 or greater to be satisfactory. It is difficult to ascertain from the literature an exact indication of a ‘satisfactorily high’ coefficient, however Purcell states: “Although Edwards implied that a coefficient of .90 was necessary when analysis was by the Cornell technique, the Goodenough technique ... was expected to result in a slightly lower coefficient, and it was implied that .85 would be sufficiently high.”

From Table 21 (p.169), it was reported that the coefficient of reproducibility for the Main Respondent group (N=310) was .87; with values for the smaller groups and university students ranging from 0.84 to 0.91. These are promising results, hence it may be concluded from the Goodenough analysis that the NMAAI has been successful in incorporating a hierarchical structure.

As far as the Binomial Test calculations were concerned (pp.169-172), the result of $Z = 3.294$ means that the high proportion of respondents’ scores displaying hierarchy (according to the definition of this study) was very unlikely to be due to chance.

It is acknowledged that small numbers of respondents were used in the trialling of the NMAAI, apart from the Main Respondent, and the Test-retest groups. Nonetheless, these exercises certainly suggest a measure of support for basing an assessment instrument on the affective taxonomy of Krathwohl *et al.*. Attention must be drawn at this point,

however, to the considerable time and attention required to develop such an instrument. This applies not only to establishing the correct taxonomic level placements for the item statements, but also to dealing with the vernacular of the adolescent group. This finding supports that of other researchers, for example, Purcell and Dave [see Chapter IV, p.98].

3. IMPLICATIONS FOR SECONDARY SCHOOL GENERAL MUSIC EDUCATION

It will be recalled from Chapter II that the purpose of the present study was:

- (1) to examine the issue of assessment in the affective domain, with particular reference to New Zealand secondary school general music education, and
- (2) to make a practical contribution to the assessment of elements in the affective domain, by addressing problems and concerns raised by the examination of the various affective domain issues.

The NMAAI, developed as part of the 'contribution to the assessment of elements in the affective domain' inquiry, was intended to fulfil a practical purpose - particularly as far as New Zealand secondary school general music education was concerned.

To this end, prior to the construction and testing of the NMAAI, six objectives were identified [Chapter II, p.51] as being desirable to incorporate into an attitude assessment instrument in music. Those objectives will now be examined in the light of the results obtained from the trialling of the NMAAI:

- (1) The instrument must "furnish details of the 'status quo' of students' affective relationships with music." (p.51) [see also discussion in Chapter II, p.41]

- (a) As the NMAAI was developed according to the principles of the affective taxonomy of Krathwohl *et al*, then theoretically the position of a respondent on the affective continuum will be locatable. By this, one is referring to the fact that the instrument reflects the ascending-categories structure of the taxonomy, with the item statements of each level corresponding to types and intensities of affective behaviors (attitudes in particular). Students' performances on the NMAAI may be directly related to the affective continuum (*as represented by* those groups/levels of item statements), and conclusions drawn from there. For example, students who respond well to item statements of levels one and two only, are not displaying the same depth of attitude to music as students responding well to *all* the levels. The former may well be showing an interest in musical activities, particularly if being directed to do so by others, but a level of *commitment* will not be as evident as with the latter.

This premise was also supported by other researchers, for example, Lewy [under the guidance of Bloom and Masia], Noble, and Colwell.

- (b) The validity trial for the NMAAI which involved the 'Music Production Students' respondent group yielded positive results; indicating the capacity of the instrument to correctly identify respondents' attitudes to music.
- (2) The instrument must not be "drawing primarily from the cognitive/intellectual abilities of a student." (p.51)
- (a) The respondent may obtain assistance from the teacher at any stage during the completing of the NMAAI.

- (b) The NMAAI is not concerned with the fact that the respondent has received any formal musical training, or has completed a school music education programme of any kind. Cognitive abilities such as factual knowledge recall are not assessed.
- (3) The instrument should reflect “the contemporary call for formative evaluation.” (p.51)
 - (a) The fact that the NMAAI is based on the affective taxonomy of Krathwohl *et al* means that the respondents are being assessed according to the inherent hierarchical structure of the model [refer also to (1)(a)]. Comparisons with the performances of other respondents are not necessary.
 - (b) The NMAAI may be used at any time during a course of study; it is not reliant on students completing a specific unit of work, for example. Hence it is not primarily designed for summative purposes.
- (4) The instrument will embrace “the framework of Krathwohl’s taxonomy of affective objectives.” (p.51) [see also discussion in Chapter II, pp.43,44]
 - (a) The NMAAI was originally developed to illustrate the appropriateness of basing an affective assessment instrument on the affective taxonomy. Hence the underlying principles of the NMAAI are those of the model.
- (5) The instrument needs to be “non-threatening for the respondent, with a view to minimising detrimental environmental factors.” (p.51) [see also discussions in Chapter II, pp.46-47 ‘credibility gap’ and p.48 ‘anonymity issue’.]
 - (a) Administering the NMAAI does not require ‘test’ conditions.

- (b) The fact that the NMAAI does not require factual or theoretical knowledge in music lessens the ‘cognitive’ anxiety for the respondents.
- (c) The NMAAI allows students to respond according to their own musical preferences; in this way, the respondents are not forced to contemplate a musical idiom that is alien to them.
- (d) The attitudes of the students involved in the trialling of the NMAAI made it clear that the instrument is non-threatening: in other words, they were quite amenable about furnishing their responses.
- (e) During the pilot testing, statistical differences between the anonymous and the named responses did not arise. The students who were not required to write their names on their responses were often enthusiastic about doing so.

A percentage of the NMAAI respondents were also anonymous. The following table illustrates the similarity of the anonymous group and the Main Respondent group results:

Table 25: NMAAI Results - A Comparison of the Anonymous and Main Respondent Groups.

Group	No.	A	B	C	D	E
Anonymous	49	33%	55%	12%	71%	6%
Main Respondent	310	33%	53%	14%	74%	11%

- (6) The instrument needs to address the issue of cultural/social unfairness. (p.51) [see also discussion in Chapter II, p.45]
- (a) The term 'music' in the NMAAI is unrestricted. It is stated at the beginning of the instrument that "The use of the word 'Music' in this inventory is very general; it can mean any style from Rock to Classical, including Maori and other ethnic music. You are to respond according to your own preferences." The respondent information sheet showed that preferred music types did indeed range widely; although the majority of secondary school students indicated preferences for popular music styles such as rock, heavy metal, rap etc.
- (b) The trialling of the NMAAI indicated that the instrument may be culturally fair [see p.185].

Hence one may conclude that the NMAAI, as an instrument for use in general music education, contains demonstrably beneficial properties. In the New Zealand setting, specific assessment objectives from the new music education syllabus for schools³⁴¹ that may be addressed by the NMAAI are 1. how involved students are with music, and 2. the breadth and depth of students' interests and how such interests may have changed [if, for example, the NMAAI is administered before and after a course of instruction]. The third affective objective stated by the new music education syllabus, *the extent to which the student has a positive, open attitude to a wide range of music*, is not catered for by the NMAAI, as students are able to respond to the Instrument according to their own specific musical preferences. A teacher might gain some indication of 'range' of attitude, however, from the respondent information sheet.

³⁴¹ Refer pp. 68-69 of this thesis.

Examples of uses to which the NMAAI may eventually be put in a teaching situation include:

- 1) Assessing the music attitudes of new classes (i.e. discovering the 'status quo') in order to apply/modify teaching programmes
- 2) Aiding in the selection of students for specific programmes (e.g. extra-curricula instrumental training)
- 3) Formative evaluation of students according to affective teaching objectives
- 4) Assessing teaching programmes (perhaps by allowing respondent anonymity, to encourage even greater reliability of results).

4. IMPLICATIONS FOR FURTHER RESEARCH

(1) The NMAAI:

There are several areas of importance in the NMAAI that one might profitably take up and research to greater lengths. To begin with, one could focus on any one of the smaller testing groups and make a full study of their significance to the application of the instrument. For example with the 'Slow Learner' group, greater numbers could be employed in order to further test the 'cognitive versus affective assessment' issue.

Similarly with the 'adult' groups, one could ask such questions as 'At what point (with respect to age of respondents), if at all, does the NMAAI become unreliable with respondents who are not musically committed?' (It will be recalled that the statistical results of the 'adults with no particular musical association' group were quite unpredictable, indicating "a lack of compatibility between these respondents and the NMAAI." (p.182))

A valuable contribution, too, would be effected if the NMAAI were to be administered to groups of students over an extended period of time, at distinct intervals. [As one of the reasons for developing the NMAAI was to "furnish details of the 'status quo' of students'

affective relationships with music", this particular exercise was beyond the scope of the present study.] Attempts then could be made to monitor attitudes and their development; an exercise which subsequently might be used as a guide, not only for student assessments, but also for evaluating the success of school music teaching programmes.

Further work, too, could be undertaken with respect to the placement of 'attitude' in the affective continuum; also regarding whether it is possible to construct items that will adequately represent 'Level 5' of the Taxonomy at the secondary school level.

(2) Assessment in the affective domain:

Apart from the further research one might undertake with regard to the NMAAI, the present study has initiated inquiry into further, wider issues. In particular, exploratory investigations were begun into the interaction of the affective and cognitive domains. Comparisons could now be made between data obtained from administering the NMAAI to groups of students of widely differing abilities, and data collected from selected cognitively-weighted instruments, such as music aptitude, and general educational tests (using the same testing groups of course.) The objective would be to support, or possibly refute, the findings of the present study with regard to the 'slow learner' group trials [see p.185].

Another area of investigation could be the effectiveness (or otherwise) of teacher-ratings, and student self-ratings, with regard to affective domain assessment objectives. The present study merely touched upon this complex issue.

Of value also to the issue of assessment in the affective domain would be further study into affective growth itself; the question being whether focussing on affective objectives in the curriculum actually influences or benefits students' affective development. Krathwohl *et al* state:

*The evidence suggests that affective behaviors develop when appropriate learning experiences are provided for students much the same as cognitive behaviors develop from appropriate learning experiences.*³⁴²

More research is needed to support this theory, however.

*** **

As many indications of a person's positive response to, or involvement with music are not 'measurable' in a traditional sense, they are in danger of escaping notice, commendation and encouragement.

Students should not 'fail' music per se. All are capable of finding a 'niche' for their relationship with music: this niche should be located and developed, otherwise it may be destined for lifelong latency. As it is stated in the New Zealand Syllabus For Schools: Music Education - Early Childhood To Form Seven:

*The aim of music education is to involve people in the active, creative processes of making and listening to music, in ways that promote individual aesthetic growth and fulfilment.*³⁴³

³⁴² D. Krathwohl, B. Bloom, and B. Masia, (1964), op. cit., p. 20.

³⁴³ Syllabus For Schools: Music Education - Early Childhood To Form Seven, New Zealand Department of Education, (1989), p. 6.

The key words, 'promote individual aesthetic growth and fulfilment', surely appeal to affective domain concerns, and until greater prominence is accorded to the assessment of such objectives in education, they will remain just that: only words.

Appendix A

THE FIVE CATEGORIES OF THE AFFECTIVE TAXONOMY OF KRATHWOHL, BLOOM AND MASIA.

1.0 Receiving (Attending)

1.1 Awareness

1.2 Willingness to receive

1.3 Controlled or selected attention

2.0 Responding

2.1 Acquiescence in responding

2.2 Willingness to respond

2.3 Satisfaction in response

3.0 Valuing

3.1 Acceptance of a value

3.2 Preference for a value

3.3 Commitment (conviction)

4.0 Organization

4.1 Conceptualization of a value

4.2 Organization of a value system

5.0 Characterization

5.1 Generalized set

5.2 Characterization

Appendix B

HEVNER INSTRUMENT

The following two scales constitute Kate Hevner's 'Tests For Attitude Toward Music' as outlined in her article "Appreciation of Music and Tests for the Appreciation of Music" (1934).

Scale A

1. I do not attend many concerts but when I do go, I always find something to enjoy.
2. Music provides one of the greatest pleasures of my life.
3. I enjoy dance music, and singing, but I am indifferent to the higher types of music.
4. Music in the life of the growing individual is neither particularly beneficial nor harmful.
5. I think music is not only a waste of time, but I despise those who devote themselves to it.
6. I am not a good judge of the merits of music because much of what I hear I do not understand.
7. I find music valuable because it lifts us out of the humdrum level of life and brings us up to a higher plane.
8. That anyone should want to devote his whole life to music seems utterly incomprehensible to me.
9. I should say that appropriate music always adds to my enjoyment of any drama, or ceremony of any sort.
10. Compared with other forms of entertainment and recreation, music has less to offer me.

11. Music is the greatest of all the arts, for not only does it provide for the greatest ecstasies of experience, but through its many forms it is capable of the widest and broadest appeal.
12. I have the greatest admiration for music and those who make it, but I do not care for much of it myself.
13. I like music but I could live the rest of my life without hearing any more of it.
14. Good music is so cheap, and so accessible and offers so much pleasure that much more time should be spent on it in the public schools.
15. I think that too much music for children makes for an unbalanced mental growth.
16. My facility for listening to music is subject to fatigue, but gives me keen pleasure for short periods.
17. If I thought that music could do for me what it seems to do for some people, I would be willing to spend more time on it.
18. Music may be an important part of the life of some people, but it has been vastly overestimated.
19. I enjoy songs, where the words and music supplement each other, but pure music does not interest me.
20. To me there are few things which could be more boresome than grand opera or symphony concerts.
21. Music is not a luxury, it is a necessity: it rests and relaxes the body and stimulates and refreshes the mind.
22. Although I do not spend a great deal of my leisure time on music I would not be willing to give up the time I do spend on it.
23. To me music is too formal and artificial: it seems to have lost its natural appeal to the tastes and feelings.

24. I am happier listening to music than at any other time.
25. I believe the world would be just as well off if there were no music in it.

Scale B

1. The aesthetic perfection of truly great music surpasses that of all other artistic forms.
2. Although I enjoy music I prefer recreations in which I can take a more active part.
3. I have a casual interest in music.
4. There is a rich and varied pleasure in music which those who have once indulged can never allow themselves to be without.
5. Although I often listen to music when I find it at hand, I do not actively seek it out for myself.
6. I am against music for I believe the study of it makes people peculiar and narrow-minded.
7. Most types of music are effeminate and have little possible interests for persons engaged in the more active pursuits of life.
8. Music seems to have established itself as a part of our present day culture and I feel it is an intellectual duty to keep in touch with it.
9. I set a high value on music because it provides entertainment and relaxation for so many people.
10. Compared with other forms of entertainment and recreation, music has less to offer me.
11. The one or two real thrills that I get out of a concert are not worth the labour of listening to the whole concert.

12. I do not care for music myself, but I am not opposed to it in any way.
13. I believe strongly in the beneficial and pleasurable effects of music, but do not care to take an active part in it myself.
14. It is too much for me to believe that the musical enthusiasm of most people is genuine and sincere.
15. I think the importance of music has been overestimated.
16. Living would be a much more dull and drab affair were it not for the beauties of music.
17. I am not a music enthusiast but I would be sorry to see it disappear from our civilisation.
18. I get less to repay me for time spent on music than for time spent on my other recreations.
19. In music I find the best expression of some of the greatest realities of life.
20. Music has nothing to offer me. It not only bores me, but it is a distinctly unpleasant experience.
21. Music gives me certain pleasures and satisfactions which I would not want to do without.
22. I believe music is an ideal accompaniment for dance or drama, but I do not care to listen to it for itself.
23. Interest in music has been fostered artificially by rich philanthropies and enthusiastic artists. It has no real value to the general public.
24. On the whole, I am indifferent to music, but occasionally I find a great deal of pleasure from some chance hearing of a musical number.
25. To me there is nothing so beautiful, and so worth while in life as good music.

Appendix C.

NOBLE INSTRUMENT

The following is from the 'attitude to music' assessment instrument developed by Robert Noble in his study "A Multivariate Analysis of Factors in the Backgrounds of Wyoming Adults Related to their Attitudinal Levels Concerning Music".

Instructions for completing the 'Attitudes Toward Music Inventory' were thus:

1. This research is most dependent upon establishing an accurate assessment of adults' choices and feelings about music and music participation.
2. There are no "right" or "wrong" answers; therefore, respond to each question as you feel and not as you might think the inventory-maker might want. Please respond to all questions; do not omit any.
3. Most of the questions can be answered by simply placing a check mark in the space to the left of the identifying letter you choose. In some, you will be asked to rate (by numbers) several possible answers. In the final questions, you will be asked to respond in your own words.

NOTE: Items 47 - 55 were not made available to the present study.

1. Are there any special kinds of music you like to listen to?
___ a) yes
___ b) no
2. Are you aware of your feelings brought about by the music you hear?
___ a) yes
___ b) no
3. Do you notice changes in loudness and softness in listening to music?
___ a) yes
___ b) no

4. If you have any children of your own involved in school music organizations, do you belong to a music-parents organization to help the school music program?
- ☐ a) yes
 - ☐ b) no
 - ☐ c) I have no children involved in school music organizations
 - ☐ d) Our school has no music-parents organization.
5. How often have you asked your local radio station to provide more classical music?
- ☐ a) I wouldn't ask them to do this
 - ☐ b) I never have
 - ☐ c) Once or twice
 - ☐ d) More than twice
6. Have you volunteered to arrange for a music program for a service club meeting or another organisation?
- ☐ a) yes
 - ☐ b) no
7. In movies and on TV, are you consciously aware of the existence of the background music to the actors' words and actions?
- ☐ a) yes
 - ☐ b) no
8. Would you like to learn how to play a musical instrument, if you had the opportunity?
- ☐ a) yes
 - ☐ b) no
9. At social or church gatherings, do you usually join in the group singing?
- ☐ a) yes
 - ☐ b) no
10. How often do you seek out and listen to music on radio?
- ☐ a) less than once a month
 - ☐ b) 1 to 3 times a month
 - ☐ c) once a week
 - ☐ d) 2 to 6 times a week
 - ☐ e) every day

11. Do you practice by yourself on any instrument or voice once a week or more often?
___ a) yes
___ b) no
12. Singing with a group is enjoyable.
___ a) I strongly agree
___ b) I somewhat agree
___ c) I am undecided
___ d) I somewhat disagree
___ e) I strongly disagree
13. Have you ever had the urge to compose a melody of your own?
___ a) yes
___ b) no
14. How often do you seek out and listen to music on records or tapes?
___ a) less than once a month
___ b) 1 to 3 times a month
___ c) once a week
___ d) 2 to 6 times a week
___ e) every day
15. Would you attend concerts of school or community groups three times a year or more?
___ a) yes
___ b) no
16. Would you be interested in hearing a mass sung in Latin in the place that you worship?
___ a) yes
___ b) no
17. Have you volunteered to join some large community music performance group (such as a jazz ensemble, vocal ensemble, community band)?
___ a) yes
___ b) no

18. Assuming you might belong to any type of musical group within your community, which of the following might you most typically do?
- ☐ a) Attend the rehearsal without being reminded by others.
 - ☐ b) Go if I am reminded
 - ☐ c) Try to get out of going.
 - ☐ d) Don't go even if I am reminded.
 - ☐ e) Don't go and nobody objects.
19. Supposing you were buying a popular music record. For which of the following reasons might you be most likely to select your record?
- ☐ a) because of the name of the individual or group who had recorded it.
 - ☐ b) because friends of mine had recommended it.
 - ☐ c) because I liked it best.
 - ☐ d) because it was the most musical recording.
20. Would you like to attend a performance of an opera?
- ☐ a) yes
 - ☐ b) no
21. While it can be used in both, one instrument from the following is commonly used in the stage or dance band but seldom in the symphony orchestra:
- ☐ a) trumpet
 - ☐ b) saxophone
 - ☐ c) drums
22. Do you like either to sing or to play a musical instrument?
- ☐ a) very much
 - ☐ b) somewhat
 - ☐ c) not at all
23. Which of the following popular groups do you like best?
- ☐ a) Chicago
 - ☐ b) Allman brothers
 - ☐ c) Gladys Knight and the Pips
 - ☐ d) Emerson, Lake, and Palmer

24. What kinds of music do you most like to listen to most of the time?
- ☐ a) Instrumental art (e.g., classical, symphonic)
 - ☐ b) Vocal art (e.g., opera), choral groups (classical performances, e.g., Norman Tabernacle Choir)
 - ☐ c) jazz
 - ☐ d) folk
 - ☐ e) rock
 - ☐ f) country-western
 - ☐ g) soul
 - ☐ h) popular ballads (e.g., barbershop, male vocalists, romantic)
 - ☐ i) blues
 - ☐ j) background music
 - ☐ k) other popular
25. In a community choir, orchestra or band, do you:
- ☐ a) volunteer to participate.
 - ☐ b) participate without being urged.
 - ☐ c) participate but with some reluctance.
 - ☐ d) try to get out of doing it.
 - ☐ e) refuse to do it.
 - ☐ f) lack interest in this type of activity.
26. Are you usually conscious of the tempo (speed, as being fast, medium, or slow) of music when you hear it?
- ☐ a) yes
 - ☐ b) no
27. What do you like best in rock music?
- ☐ a) the words
 - ☐ b) the beat
 - ☐ c) the melody
 - ☐ d) the harmony
 - ☐ e) the types of voices

28. What kinds of music do you most like to sing or play on an instrument?
- ☐ a) art music (e.g., opera, other classical)
 - ☐ b) jazz
 - ☐ c) folk
 - ☐ d) rock
 - ☐ e) country-western
 - ☐ f) soul
 - ☐ g) popular ballads
 - ☐ h) blues
 - ☐ i) other popular
29. In attending concerts or listening to recordings, what do you primarily listen for?
- ☐ a) whether I know the composition well enough to hum the melodies.
 - ☐ b) analyzation of the musical structure (form) of each composition
 - ☐ c) interrelationships of the musical elements
 - ☐ d) whether it is well done or not (in tune, good attacks, well-balanced, etc.)
 - ☐ e) the pleasure or dissatisfaction I derive from the total effort.
30. When the Saturday afternoon Metropolitan Opera broadcasts come on, which of the following do you normally do?
- ☐ a) Leave the broadcast of the opera on.
 - ☐ b) Change the dial to another station.
31. From the following choices, which one would come closest to representing what you do concerning school music concerts?
- ☐ a) I'm not interested enough to attend.
 - ☐ b) I attend if with a group of friends.
 - ☐ c) I attend if I have a son or daughter performing in the group.
 - ☐ d) I attend them when I can, if it is one I think I'd like.
 - ☐ e) I look forward to attending them.
32. Which of the following would you see as being the most appropriate medium for performing the Star Spangled banner prior to a baseball game?
- ☐ a) a jazz band.
 - ☐ b) a symphony orchestra.
 - ☐ c) a concert band.
 - ☐ d) an organ.

33. What approximate percentage of the recordings you have would fall into each of the following categories?
- ☐ a) instrumental or vocal art (classical, symphonic, operatic).
 - ☐ b) jazz.
 - ☐ c) folk.
 - ☐ d) rock.
 - ☐ e) country-western.
 - ☐ f) soul.
 - ☐ g) popular ballads.
 - ☐ h) blues.
 - ☐ i) background music.
 - ☐ j) other popular.
34. What percentage of your expendable income do you see as ideal to spend on live musical performances?
- ☐ a) less than 5 percent.
 - ☐ b) 5 percent to 15 percent.
 - ☐ c) 15 percent to 25 percent.
 - ☐ d) more than 25 percent.
35. Which, among the following, is the closest to your attitude concerning current pop music?
- ☐ a) It is the most important type of music. It is vital, alive and reflects what people think and feel today.
 - ☐ b) I'm willing to go along with it. I don't feel any strong liking or dislike for popular music.
 - ☐ c) Popular music is both good and bad. All the ingredients of the best of any kind of music are present in the best of popular music.
 - ☐ d) Popular music is okay for kids, but it is not now appealing to me.
 - ☐ e) Popular music turns me off. I don't like to listen to it.
36. Music compositions to be performed by electronic means have been developed in recent years. Do you think that electronic music can be "good" music?
- ☐ a) yes
 - ☐ b) no

37. Which of the following might come closest to your visualization of the chief purpose of music education in your schools?
- ☐ a) The development of excellence in musical performance.
 - ☐ b) The involvement of a majority (if not all) of students in some musical activity.
 - ☐ c) The development, to the fullest extent possible, of the performance excellence of students with the highest potentials in music.
 - ☐ d) The development of aesthetic sensitivity, through musical understanding, of the majority of students as consumers of music.
38. Professional performing groups, both pop and classic, are integral to the maintenance and development of our musical culture. Because of the financial difficulties almost all of them are presently having in this country, they should be directly subsidized by the Federal government, as such groups are in Europe.
- ☐ a) I disagree
 - ☐ b) I agree
39. Have you organized some musical activity such as a dance band, rock group, small vocal or instrumental group, or music-study group?
- ☐ a) yes
 - ☐ b) no
40. Rate the following in order of importance to you. Assuming the following television shows were all on at the same time, which would be your order of priority in watching?
- ☐ a) The World of Disney
 - ☐ b) Kojak
 - ☐ c) Evening at the Symphony (Boston Pops Orchestra)
 - ☐ d) Masterpiece Theatre
 - ☐ e) Cher
41. Most of the music composed by the composers of the 18th and 19th centuries is superior to most of the serious classic music composed today.
- ☐ a) I agree
 - ☐ b) I disagree

42. A number of subject areas (such as mathematics/science, humanities, fine arts, vocational subjects, work experience, and physical education) normally make up the course offerings of our public high schools. Which one of the following would be closest to your viewpoint concerning the balance of these offerings for every student?
- ☐ a) The high school should emphasize entry skills into specific vocations. Therefore, students should be able to select a majority of their classes aimed at their chosen specific vocation.
 - ☐ b) We are now in a mathematics/science oriented world. We should de-emphasize other areas and place chief emphasis on mathematics and science offerings.
 - ☐ c) There is not enough emphasis today in humanities and Fine Arts areas. These areas need strengthening in our schools.
 - ☐ d) Fine Arts should be the central core of the curriculum for all students with other areas supplied to fit the other needs of individual students.
43. Of the following, rate (1-5) in order of importance to you, music's most important role in contemporary life.
- ☐ a) Release from tensions of every day life.
 - ☐ b) A means of human understanding.
 - ☐ c) A form of human communication not dependent on words.
 - ☐ d) A realization of the highest in human Art forms.
 - ☐ e) A source of self expression.
44. Among your personal books, do ten or more concern some aspect of music or composers?
- ☐ a) yes
 - ☐ b) no
45. What percent of your expendable income do you see as ideal to spend on live musical performances?
- ☐ a) less than 5 percent.
 - ☐ b) 5 percent to 15 percent.
 - ☐ c) 15 percent to 25 percent.
 - ☐ d) more than 25 percent.

46. Have you played in a dance band, adult community band or orchestra, or sung in a community or church choir for a period of 5 years or more?

☐ a) yes

☐ b) no

Appendix D

ITEM LEVEL PLACEMENT FOR NOBLE INSTRUMENT

LEVEL (Noble)	LEVEL (Taxonomy)	ITEMS							
1	1.1	1	2	3	7	26			
2	1.2	8	15	16	20	30			
3	1.3	19	21	23	27	32			
4	2.1	4	9	13	18	25			
5	2.2	6	10	11	14	17			
6	2.3	5	12	22	24	28			
7	3.1	35	36	41	48	52			
8	3.2	29	31	33	34	45			
9	3.3	39	44	46	47	51			
10	4.1	37	38	42	49	50			
11	4.2	40	43	53	54(1)	54(2)	55(1)	55(2)	

Appendix E

LEWY INSTRUMENT

The following is the music measurement instrument developed by Arie Lewy in his PhD study “The Empirical Validity of Major Properties of a Taxonomy of Affective Educational Objectives”.

For statements 1 - 42, 50 - 66, 69 - 77, and 79 - 82 the respondent has the options of ‘true’ and ‘not true’. For statements 43 - 49 the respondent has the options of ‘yes’ and ‘no’.

The statement “Do not forget that in this questionnaire the word music stands for classical music only” is repeated on alternate pages of the instrument.

1. I have never had the desire to go to a concert.
2. It is very unusual for me to play the radio for the purpose of listening to music. (Do not forget what the word music stands for in this questionnaire!)
3. If my parents or friends play the radio and listen to music, I prefer to leave the room.
4. I have never had the wish to improve my understanding of music.
5. I would like to devote some time to listening to good music.
6. Going to a concert is for me a more attractive entertainment than going to a movie.
7. Being with friends, or in a party, I usually welcome suggestions to listen to records of classical music.
8. Sometimes when I am tired, it “refreshes” me to listen to some short musical piece.
9. There are several musical pieces, which I like so much that I would like to hear them performed by various musicians.

10. It is unusual for me to listen to music continuously for a period of two hours.
11. I do attend concerts quite regularly (on the average at least once in six weeks).
12. I do attend concerts at least once in a year.
13. If certain musical pieces are on the program on radio, I cannot resist and I have to hear them.
14. It is very unusual for me to plan in advance which musical program I shall listen to on the radio in the coming week.
15. When listening to music, I prefer to concentrate on the music and not do anything else, such as reading.
16. When I attend a concert, it is usually at the urging of somebody else (my parents or friends).
17. I usually do not remember the titles of musical pieces which I have listened to.
18. Sometimes I like to follow a musical performance from a printed 'score'.
19. I would not be willing to give up the time I do spend on listening to music.
20. Listening to good music is a part of my everyday life.
21. Music bores me.
22. Compared with other forms of entertainment and recreation, music has less to offer me.
23. Listening to long musical pieces makes me usually tired.
24. I do enjoy modern symphonic music.
25. Non-melodic sections of musical compositions do not appeal to me.
26. I try to memorize melodies from music which I hear.

27. Coming across familiar melodies when listening to music causes me great pleasure.
28. When listening to music, I usually try to make judgements about the quality of the performance.
29. To me there are few things which could be more boresome like grand operas and symphony concerts.
30. Music often provides me with entertainment and relaxation.
31. It annoys me to be disturbed when I am listening to music.
32. Going to a concert seems to be more of a social event than a musical experience.
33. Music provides me with one of my greatest pleasures.
34. To me music is just noise.
35. It is not unusual for me to compare performance of the same music by different musicians.
36. If I have been impressed by a musical piece, I usually look for additional opportunities to hear it again.
37. When buying records of classical music I do pay attention to who is the performer of the music.
38. I like to exchange views on the quality of performance with others who have heard the same program.
39. In forming an opinion on performance, I consider the technique of the performers.
40. Listening to a composition, I do not care too much about what historical period it belongs to.
41. In listening to a musical piece which I have never heard before, I try to form an opinion about the nature of this composition.

42. Music is for me a continuous flow of rhythm lacking any form of structural organization.

Have you ever tried to explain or discuss the following topics?

43. What makes a musical composition immortal?
44. How music can create or express feelings.
45. What are the differences between the musical styles of various historical periods?
46. How a particular type of music reveals something about the people who created it?

Many literary themes have musical and drama versions, e.g. Faust, etc.

47. Were you ever curious to hear the musical version of a drama which you liked?
48. Could you give some more examples of drama which have some kind of musical version too?
49. Have you ever tried to compare the merits and demerits of the drama version and musical version of the same theme?
50. I find music valuable because it lifts us out of the monotonous level of life and takes us to a higher plane.
51. Music, through the variety of its forms, is capable of the widest and broadest appeal.
52. The aesthetic perfection of truly great music surpasses that of all other artistic forms.
53. Music make people narrow-minded.
54. Most types of music are effeminate and have little possible interest for persons engaged in more active pursuits of life.
55. I would be sorry if music began to disappear from our civilization.

56. Love for music can create a feeling of understanding among people of various countries.
57. The cultural value of music is not any less than that of literature.
58. One does not have a well-rounded general education if he is not familiar with the best musical compositions.
59. One of the shortcomings of current TV programs is the relatively low proportion of good music among the programs.
60. In my eyes, classical music does not represent a higher level of art than popular music.
61. I would like to read some interesting books on music.
62. It is not unusual for me to read criticisms on musical pieces which I had heard.
63. I often read articles which deal with recently composed music.
64. It is not unusual for me to read newspaper reviews of musical performances which I attended.
65. I usually read newspaper reviews of new recordings.
66. I like to read about local musical events.
67. I would like to know more about musical theory.
___ A. Yes
___ B. No, it does not interest me.
___ C. No, I have sufficient knowledge.
68. I would like to know more about modern music.
___ A. Yes
___ B. No, it does not interest me.
___ C. No, I have sufficient knowledge.
69. I would like to know more about the history of music.

70. I would like to become more familiar with famous musical pieces.
71. I would like to know more about differences between styles in music.
72. It is interesting for me to hear remarks on a musical performance which I happened to hear.
73. I like to listen to conversation on topics related to music.
74. It is not unusual for me to talk about topics related to music.
75. If I like a certain musical piece, I often recommend it to my friends for the purpose of listening to it.
76. In several instances I tried to convince my friends that they should devote some time for listening to music.
77. I like to listen to records which contain explanations or discussion of the music played.
78. I would like to start to build a collection of records of serious music. (Mark one of the answers.)
- ___ A. Yes
- ___ B. No, I am not interested
- ___ C. I have already

You are presented here a page of a daily newspaper. Read the passages which are marked, and put mark in proper place.

79. I have never paid attention to articles like this.
80. This seems to me an interesting article.
81. I used to read [sic] at least the titles of articles like this even if I do not read the whole article.
82. Quite often I read articles like this with great interest.

83. Leonard Bernstein is a

- ☐ A. composer
- ☐ B. violinist
- ☐ C. tenor
- ☐ D. cellist

84. Toscanini is the name of a

- ☐ A. famous opera
- ☐ B. conductor
- ☐ C. pianist
- ☐ D. place of musical festivals

85. Marian Anderson is [a]

- ☐ A. singer
- ☐ B. violinist
- ☐ C. cellist
- ☐ D. organist

86. [A] Famous composer of operas was

- ☐ A. Schubert
- ☐ B. F. Liszt
- ☐ C. Mendelssohn
- ☐ D. Verdi

Appendix F

ITEM LEVEL PLACEMENT FOR LEWY INSTRUMENT

Level	Item																
1	1	2	3	4	5	8	21	34	42	47	48	53	54	61	67	68	
	69	70	71	79	80	81	83	84	85	86							
2	7	9	10	12	16	23	24	25	27	29	30	32	36	72	73	78	
3	6	11	13	14	15	17	18	19	20	22	26	31	33	37	40	41	51
	52	55	56	57	58	59	60	62	63	64	65	66	74	75	76	77	82
4	28	35	38	39	43	44	45	46	49	50							

Appendix G

ADAPTED LEWY INSTRUMENT

The following is the adapted version of the Lewy Instrument (see Appendix F) as used in the present study.

For statements 1 - 46 and 51 - 54 the respondent has the options of 'true' and 'not true'. For statements 47 - 50 the respondent has the options of 'yes' and 'no'.

1. I have never had the desire to go to a concert.
2. It is very unusual for me to play the radio for the purpose of listening to music.
3. Sometimes when I am tired, it "refreshes" me to listen to some music for a period of time.
4. There is some music which I like so much that I would like to hear it performed by various musicians.
5. When I attend a concert or other live musical performance, it is usually at the urging of somebody else (my parents or friends).
6. I would like to improve my understanding of music.
7. I would like to devote some more time to listening to music.
8. Being with friends, or at a party, I usually welcome suggestions to listen to music.
9. Listening to good music is a part of my everyday life.
10. Music bores me.
11. If I have been impressed by a musical piece, song etc., I usually look for additional opportunities to hear it again.

12. Compared with other forms of entertainment and recreation, music has less to offer me.
13. I do attend concerts or other live musical performances at least once in a year.
14. Listening to music for long periods of time usually makes me tired.
15. It is quite usual for me to listen to music continuously for a period of two hours.
16. Going to a concert or other live musical performance is for me a more attractive entertainment than going to a movie.
17. Love for music can create a feeling of understanding among people of various countries.
18. Music often provides me with entertainment and relaxation.
19. When listening to music, I prefer to concentrate on the music and not do anything else, such as reading.
20. I would not be willing to give up the time I do spend on listening to music.
21. I try to memorize melodies from music which I hear.
22. When listening to music, I usually try to make judgements about the quality of the performance.
23. I like to read about local musical events.
24. Going to a concert seems to be more of a social event than a musical experience.
25. To me, music is just noise.
26. I often read articles which deal with recently composed music.
27. It is usual for me to compare performance of the same music by different musicians.
28. Coming across familiar music - songs etc - causes me great pleasure.

29. I like to exchange views on the quality of performance with others who have heard the same program/record.
30. In forming an opinion on performance, I consider the technique of the performers.
31. It annoys me to be disturbed when I am listening to music.
32. I would be sorry if music began to disappear from our civilization.
33. I like to discuss with my friends topics related to music.
34. If I like a certain musical piece/song etc., I often recommend it to my friends.
35. Music provides me with one of my greatest pleasures.
36. Music, through the variety of its forms, is capable of widest appeal.
37. I would like to know more about the history of music.
38. It is interesting for me to hear remarks on a musical performance which I happened to hear.
39. The cultural value of music is not any less than that of literature.
40. I like to listen to conversation on topics related to music.
41. I like to read newspaper reviews of musical performances which I attended.
42. I usually read newspaper reviews of new recordings.
43. I would like to start to build a/(increase my) collection of records/cassettes/compact discs.
44. I do attend concerts or other live musical performances quite regularly (on the average at least once in six weeks).

45. It is usual for me to plan in advance any musical program I shall listen to on the radio or television.

46. I would like to read some interesting books on music.

Have you ever tried to explain or discuss the following topics?

47. What makes a musical composition/song etc. a classic?

48. How music can create or express feelings.

49. What are the differences between the musical styles of various historical periods?

50. How a particular type of music reveals something about the people who created it?

51. I would like to know more about musical theory.

52. I would like to know more about differences between styles in music.

53. I find music valuable because it lifts us out of the monotonous level of life and takes us to a higher plane.

54. I like to listen to records which contain explanations or discussion of the music played.

Appendix H

ITEM LEVEL PLACEMENT FOR ADAPTED LEWY INSTRUMENT

Level	Item																
1	1	2	3	6	7	10	25	37	46	51	52						
2	4	5	8	11	13	14	15	18	24	28	36	38	40	43			
3	9	12	16	17	19	20	21	23	26	31	32	33	34	35	39	41	42
	44	45	54														
4	22	27	29	30	47	48	49	50	53								

Appendix I

THE NMAAI

The following is the 'New Music Attitude Assessment Instrument' as developed in the present study.

MUSIC ATTITUDE INVENTORY

Please read the following statements, and place a tick beside the response of your choice.

The use of the word 'Music' in this inventory is very general; it can mean any style from Rock to Classical, including Maori and other ethnic music. You are to respond according to your own preferences.

There are no right or wrong answers.

NAMEAGE

- 1) Sometimes when I'm tired, it "refreshes" me to listen to some music.
 ___ a) True
 ___ b) Not true

- 2) I would like to improve my understanding of music.
 ___ a) True
 ___ b) Not true

- 3) Listening to music that I think is good is a part of my everyday life.
 ___ a) True
 ___ b) Not true

- 4) If I have been impressed by a musical piece, song etc., I usually look for other opportunities to hear it again.
 ___ a) True
 ___ b) Not true

- 5) I am usually aware of changes in loudness and softness in music.

- ☐ a) True
☐ b) Not true
- 6) I sometimes consider how a particular type of music reveals something about the people who created it.
☐ a) True
☐ b) Not true
- 7) I go to concerts or other live musical performances at least once in a year.
☐ a) True
☐ b) Not true
- 8) When listening to music, I usually try to make a judgement about the quality of the performance.
☐ a) True
☐ b) Not true
- 9) I often read articles which deal with recently composed music.
☐ a) True
☐ b) Not true
- 10) I discuss the quality of performance with others who have heard the same programme/record.
☐ a) True
☐ b) Not true
- 11) I would like to know more about differences between styles in music.
☐ a) True
☐ b) Not true
- 12) I try to assess the elements that make a musical composition/song etc. a classic.
☐ a) True
☐ b) Not true

- 13) I enjoy discussing with my friends topics related to music.
___ a) True
___ b) Not true
- 14) If I like a certain musical piece/song etc., I often recommend it to my friends.
___ a) True
___ b) Not true
- 15) I would like to know more about the music I listen to.
___ a) True
___ b) Not true
- 16) I usually read newspaper reviews of new recordings.
___ a) True
___ b) Not true
- 17) I would like to spend more time listening to music.
___ a) True
___ b) Not true
- 18) I attend concerts or other live musical performances quite regularly (on the average at least once in six weeks).
___ a) True
___ b) Not true
- 19) When I listen to a record, I like to read the notes that accompany it.
___ a) True
___ b) Not true
- 20) In forming an opinion on performance, I am critical of the technique of the performers.
___ a) True
___ b) Not true

- 21) I would like to know more about writing music down.
___ a) True
___ b) Not true
- 22) Music provides me with entertainment and relaxation.
___ a) True
___ b) Not true
- 23) To offer a complete education, a school needs to include music.
___ a) True
___ b) Not true
- 24) I like to compare and evaluate performances of the same music by different musicians.
___ a) True
___ b) Not true
- 25) I am interested in comparing musical styles from different historical periods.
___ a) True
___ b) Not true
- 26) There are certainly special kinds of music I like to listen to.
___ a) True
___ b) Not true
- 27) It is usual for me to plan in advance any musical programme I shall listen to on the radio or television.
___ a) True
___ b) Not true
- 28) I am often aware of my feelings when I listen to music.
___ a) True
___ b) Not true

- 29) I would like to learn how to play a/another musical instrument, if I had the opportunity.
- ☐ a) True
- ☐ b) Not true
- 30) I am interested in the way that music can generate or express feelings.
- ☐ a) True
- ☐ b) Not true
- 31) I have never written, or had the urge to write, a song/piece of music of my own.
- ☐ a) True
- ☐ b) Not true
- 32) At social gatherings, I don't usually join in the group singing.
- ☐ a) True
- ☐ b) Not true
- 33) In my opinion, everyone should take some music at school.
- ☐ a) True
- ☐ b) Not true
- 34) In movies, on television, and on video, I am usually aware of the background music.
- ☐ a) True
- ☐ b) Not true
- 35) Professional performing groups, either pop or classic, are important enough for me to spend my own money on tickets to attend their concerts.
- ☐ a) True
- ☐ b) Not true
- 36) Although I will listen to music when it happens to be playing, I do not actively seek it out for myself.
- ☐ a) True
- ☐ b) Not true

37) Playing an instrument or singing is a good way for a person to spend leisure time.

☐ a) True

☐ b) Not true

38) Music brings people together and fosters school and community spirit.

☐ a) True

☐ b) Not true

39) I am involved with music only when I have to be.

☐ a) True

☐ b) Not true

40) I enjoy associating with others in some musical activity.

☐ a) True

☐ b) Not true

Appendix J

ITEM LEVEL PLACEMENT FOR THE NMAAI

LEVEL	ITEMS
1	2 5 11 15 17 21 26 28 29 34
2	1 4 7 22 31 32 36 37 39 40
3	3 9 13 14 16 18 19 23 27 38
4	6 8 10 12 20 24 25 30 33 35

Note: Items 31, 32, 36, and 39 need to be marked as 'not true' by the respondent to indicate a positive reaction to the statement. For all other items, the positive response is the 'true' option.

Appendix K

RESPONDENT INFORMATION SHEET FOR THE NMAAI

The following sheet accompanied each NMAAI, to be completed by the respondent:

RESPONDENT BACKGROUND INFORMATION

(for statistical purposes only)

I. Can you play a musical instrument?

- ___ no
- ___ yes

If yes, which one/s, and for how many years have you been able to do so?

instrument	no. of years
_____	_____
_____	_____
_____	_____

II. The ethnic background to which you identify is: (please state, e.g. Caucasian, Maori, Pacific Island etc.)

III. Your preferred types of music are: (please describe, e.g. pop, classical, rock, country & western etc.)

Appendix L

SUMMARIES FOR INDIVIDUAL NMAAI ITEMS: Test-Retest Group

	ITEM NO.	FIRST TRIAL		SECOND TRIAL		TOTAL		%	
		YES	NO	YES	NO	YES	NO	YES	NO
LEVEL ONE	2	46	46	49	43	95	89	52	48
	5	75	17	76	16	151	33	82	18
	11	32	60	29	63	61	123	33	67
	15	61	31	58	34	119	65	65	35
	17	76	16	77	15	153	31	83	17
	21	34	58	38	54	72	112	39	61
	26	87	5	81	11	168	16	91	9
	28	66	26	68	24	134	50	73	27
	29	68	24	65	27	133	51	72	28
	34	77	15	82	10	159	25	86	14
LEVEL TWO	1	79	13	79	13	158	26	86	14
	4	87	5	84	8	171	13	93	7
	7	36	56	36	56	72	112	39	61
	22	87	5	85	7	172	12	93	7
	31	42	50	37	55	79	105	43	57
	32	63	29	72	20	135	49	73	27
	36	55	37	56	36	111	73	60	40
	37	78	14	78	14	156	28	85	15
	39	69	23	72	20	141	43	77	23
	40	77	15	75	17	152	32	83	17
LEVEL THREE	3	78	14	79	13	157	27	85	15
	9	38	54	46	46	84	100	46	54
	13	73	19	69	23	142	42	77	23
	14	78	14	82	10	160	24	87	13
	16	37	55	38	54	75	109	41	59
	18	6	86	8	84	14	170	8	92
	19	39	53	40	52	79	105	43	57
	23	78	14	78	14	156	28	85	15
	27	52	40	64	28	116	68	63	37
	38	75	17	79	13	154	30	84	16
LEVEL FOUR	6	56	36	44	48	100	84	54	46
	8	36	56	38	54	74	110	40	60
	10	60	32	61	31	121	63	66	34
	12	14	78	16	76	30	154	16	84
	20	24	68	24	68	48	136	26	74
	24	48	44	42	50	90	94	49	51
	25	22	70	28	64	50	134	27	73
	30	63	29	60	32	123	61	67	33
	33	58	34	66	26	124	60	67	33
	35	44	48	47	45	91	93	49	51

BIBLIOGRAPHY

- ABBS, Peter. (1987). Living Powers: The Arts in Education, The Falmer Press, London.
- ABELES, Harold F., HOFFER, Charles, R., and KLOTMAN, Robert H. (1984). Foundations of Music Education. Shirmer Books, New York.
- ADVISORY COUNCIL ON EDUCATIONAL PLANNING. (1980). "In Search of a Statement of Aims for New Zealand Education" in Schools in New Zealand Society - A Book of Readings. Eds: Graham H. Robinson and Brian T. O'Rourke. Longman Paul Ltd, Auckland.
- AHMANN, J. Stanley, and GLOCK, Marvin D. (1967). Evaluating Pupil Growth - Principles of Tests and Measurement. 3rd ed. Allyn and Bacon, Boston.
- ALLEN, Mary J. and YEN, Wendy M. (1979). Introduction to Measurement Theory. Wadsworth, Belmont, California.
- AMERICAN PSYCHOLOGICAL ASSOCIATION INC. (1974). Standards for Educational and Psychological Tests.
- ANASTASI, Anne. (1968). Psychological Testing. 3rd ed. Macmillan Company, New York.
- ANDERSON, S. B., BALL, S., MURPHY, R. T. and Associates. (1975). Encyclopedia of Educational Evaluation. 1st ed. Jossey-Bass Publishers, San Francisco.
- ANDERSON, William M. and CAMPBELL, Patricia Shehan (Eds.). (1989). Multicultural Perspectives in Music Education. Music Educators National Conference.
- ANDREWS, Francis, M. (1971). Junior High School Music. Prentice-Hall, Englewood Cliffs, NJ.
- ARONOFF, Francis Webber. (1969). Music and Young Children. Holt, Rinehart and Winston, New York.
- ASMUS, Edward P. Jr. (1985). "The Development of a Multidimensional Instrument for the Measurement of Affective Responses to Music". Psychology of Music. Vol.13, No. 1, 1985, 19-30.
- ASMUS, Edward P. Jr. (1980). "Empirical Testing of an Affective Learning Paradigm". Journal of Research in Music Education. Vol. 28, No. 3, 143-154.
- ASMUS, Edward P. Jr. (1980). "The Operational Characteristic of Adjectives as Descriptors of Musical Affect, Vols I-II". Dissertation Abstracts International. Vol. 40, No. 08, Feb 1980, 4289A-4290A.
- ASMUS, Edward P. Jr. (1981). "Course Entry Affect and its Relationship to Course Grade in Music Education and Music Therapy Classes". Journal of Research in Music Education. Vol. 29, No. 4, 257-264.
- AXELSEN, Doris. (1981). "Swedish Adolescents and the Phonogram" in Stocktaking of Musical Life. Ed: Desmond Mark. ISME Edition No. 1.
- BALL, Samuel (Ed). (1977). Motivation in Education. Academic Press, New York.

- BARNES, Stephen H. (Ed). (1982). A Cross-Section of Research in Music Education. University Press of America, Washington DC.
- BENSON, Jeri, and HOCEVAR, Dennis. (1985). "The Impact of Item Phrasing on the Validity of Attitude Scales for Elementary School Children". Journal of Educational Measurement. Vol. 22, No. 3, Fall 1985, 231-240.
- BENTLEY, Arnold (1975). Music in Education - A Point of View. NFER Publishing Company, Berks.
- BESSOM, Malcolm E., TATARUNIS, Alphonse M., and FORCUCCI, Samuel L. (1974). Teaching Music In Today's Secondary Schools. Holt, Rinehart and Winston, New York.
- BIRNEY, R. (1964). "The Effects of Grades on Students". The Journal of Higher Education. 35/2, Feb 1964, 96-98.
- BLOOM, B.S., HASTINGS, J.T., and MADDAUS, G.F. (1971). Handbook on Formative and Summative Evaluation of Student Learning. McGraw-Hill, New York.
- BLOOM, B. S., MADDAUS, G.F. and HASTINGS, J.T. (1981). Evaluation To Improve Learning. McGraw-Hill, New York.
- BLOOM, Benjamin, S. (1968). "Learning for Mastery". Evaluation Comment. May 1968, Vol. No. 2, 1-12.
- BOSWELL, Jacquelyn. (1991). "Comparisons of Attitudinal Assessments in Middle and Junior High School General Music". Bulletin of the Council for Research in Music Education. No. 108, Spring 1991, pp.49-57.
- BOYLE, J. David, and RADOCY, Rudolf E. (1987). Measurement and Evaluation of Musical Experiences. Schirmer Books, New York.
- BRAKE, Mike. (1980). The Sociology of Youth Culture and Youth Subcultures. Routledge and Kegan Paul, London.
- BRERETON, J. L. (1944). The Case for Examinations. An Account of Their Place in Education with Some Proposals for their Reform. Cambridge University Press, London.
- BROCKLEHURST, J. Brian. (1971). Response to Music: Principles of Music Education. Routledge and Kegan Paul, London.
- BROWN, Donald J. (1970). Appraisal Procedures in the Secondary Schools. Prentice-Hall, Englewood Cliffs, NJ.
- BUCKTON, Roger and MANINS, Stuart. (1987). "Optimal Ages and Stages in Developing Musical Activities and Concepts, Affective Response to Music, and Instrument and Vocal Training - A Survey of Research Literature". The Canterbury Series: Studies in Music Education. No.2. Ed.: David Sell. University of Canterbury.
- BULLOCK, William J. (1971). Construction and Evaluation of a Test of Musico-Aesthetic Attitude. Unpublished Ph.D. Thesis, The Florida State University.
- BULLOCK, William J. (1973). "A Review of Measures of Musico-Aesthetic Attitude". Journal of Research in Music Education. Winter, 1973, Vol. 21, No. 4, 331-345.

- BURKE, R. S. and GRINDER, R.E. (1966). "Personality-Oriented Themes and Listening Patterns in Teenage Music and their Relation to Certain Academic and Peer Variables". The School Review. Vol. 74, No. 2, 196-211.
- BUXTON, Claude, E. (1973). Adolescents in School. Yale University.
- CADY, Henry L. (1987). "The Problem of the Rational in Music". Bulletin of the Council for Research in Music Education. No. 90, Winter 1987, 30-40.
- CARLE, Harold Ward. (1985). Ethical Implications of Aesthetics. Unpublished Ph.D. Thesis, The American University, Washington D.C.
- CARY, Sylvia Estes. (1981). Individualized Music Instruction - Traditional Music Instruction: Relationships of Music Achievement, Music Performance, Music Attitude, Music Aptitude, and Reading in Classes of Fifth Grade Students. Unpublished Ph.D. Thesis, University of Oregon.
- CHASE, Clinton I. (1974) Measurement For Educational Evaluation. Addison-Wesley Publishing Co., USA.
- CHURCHLEY, F. (1969). Music Curriculum and Instruction. McGraw-Hill Co. of Canada Ltd.
- CODD, J.A. and HERMANSSON, G.L. (Eds). (1976). Directions in New Zealand Secondary Education. Auckland.
- COHEN, Louis and HOLLIDAY, Michael. (1982). Statistics for Social Scientists. Harper and Row, London.
- COLEMAN, James S. (1961). The Adolescent Society: The Social Life of the Teenager and Its Impact On Education. The Free Press of Glencoe, New York.
- COLWELL, Richard. (1970). The Evaluation of Music Teaching and Learning. Prentice-Hall, Englewood Cliffs, NJ.
- COLWELL, Richard. (Ed). (1982). Symposium in Music Education: A Festschrift for Charles Leonhard. University of Illinois, Urbana-Champaign.
- CRONBACH, Lee J. (1984). Essentials of Psychological Testing. 4th ed. Harper and Row, Publishers.
- DAVE, R. H. (1969). "Taxonomy of Educational Objectives and Achievement Testing" in Developments in Educational Testing Vol. 1. Ed: Karlheinz Ingenkamp. University of London Press Ltd.
- DE VRIES, Bart. (1991). "Assessment of the Affective Response to Music With Clynés Sentograph". Psychology of Music. 1991, Vol. 19 no. 1, pp.46-64.
- DEWEY, John. (1897). "The Aesthetic Element in Education". Addresses and Proceedings of the National Education Association. National Education Association.
- DOWLING, W. Jay and HARWOOD, Dane L. (1986). Music Cognition. Academic Press.
- EBEL, Robert L. (1972). Essentials of Educational Measurement. Prentice-Hall, Englewood Cliffs, NJ.

- EDWARDS, Allen L. (1957). Techniques of Attitude Scale Construction. Appleton-Century-Crofts, New York.
- EDWARDS, Allen L. and KENNEY, Kathryn Claire. (1946). "A Comparison of the Thurstone and Likert Techniques of Attitude Scale Construction". Journal of Applied Psychology. 1946, 30, 72-83.
- EDWARDS, J. S. and EDWARDS, M.C. (1971). "A Scale to Measure Attitudes Towards Music". Journal of Research in Music Education. Vol. XIX, Summer, 1971, No. 2, 228-233.
- EISS, Albert, F. and HARBECK, M.B. (1969). Behavioral Objectives in the Affective Domain. National Science Teachers Association, Washington, DC.
- ELKIN, F. and WESTLEY, W. (1955). "The Myth of Adolescent Culture". American Sociological Review. XX, 1955, 680-684.
- FARNSWORTH, Paul R. (1935). "Are 'Music Capacity' Tests More Important Than 'Intelligence' Tests in the Prediction of the Several Types of Music Grades?". Journal of Applied Psychology. Vol. XIX, 1935, 347-350.
- FARNSWORTH, Paul R. (1949). "Rating Scales for Musical Interests". The Journal of Psychology. 1949, 28, 245-253.
- FARNSWORTH, Paul R. (1969). The Social Psychology of Music. 2nd ed. AMES: Iowa State University Press.
- FERGUSON, Leonard W. (1941). "A Study of the Likert Technique of Attitude Scale Construction". Journal of Social Psychology. 1941, 13, 51-57.
- FISHER, N. G. (c. 1945). "Music and the Curriculum of the Future" in Musical Education. Ed: H. Watkins Shaw. Hinrichsen, London.
- FLETCHER, Peter. (1989). Education and Music. Oxford University Press, Oxford.
- FOSTER, John. (1971). Recording Individual Progress. MacMillan Education.
- FRANKLIN, Erik. (1972). Music Education, Psychology and Method. George Harrap and Co., London.
- FREEDMAN, A. (1972). Implications with Suggested Applications of Selected Educational Theories and Taxonomic Models for the Development of Productive Thinking in the General Music Class in the Secondary School. Unpublished Mus.Ed.D.Thesis, Indiana University.
- FRITH, Simon. (1981). Sound Effects: Youth, Leisure, and the Politics of Rock'n'Roll. Pantheon Books, New York.
- FROELICH, Hildegard. (1980). "Raymond Charles Wifler: The development and application of an instrument to identify music-affect shift resulting from a college music appreciation course: A Review". Bulletin of the Council for Research in Music Education. No. 64, Fall 1980, 48-54.
- GATES, J. Terry, (1988). Music Education in the United States: Contemporary Issues. University of Alabama Press, Tuscaloosa, Alabama.

- GIBLIN, Paul T. (1981). "Affective Development in Children, An Equilibrium Model." Genetic Psychology Monographs.
- GLENN, N.E., MCBRIDE, W.B., and WILSON, G.H. (1970). Secondary School Music: Philosophy, Theory and Practice. Prentice-Hall, Englewood Cliffs, NJ.
- GOODMAN, Paul. (1971). Compulsory Miseducation. Penguin Books, [Harmondsworth] (Reprinted 1972). Copyright Paul Goodman, 1962, 1964.
- GORDON, Edwin. (1965). Manual: Musical Aptitude Profile. Houghton Mifflin Company, Boston.
- GORDON, Edwin. (1967). "A Three Year Study of the Musical Aptitude Profile". Studies in the Psychology of Music. Vol. 5. University of Iowa Press.
- GORDON, Edwin. (1971). The Psychology of Music Teaching. Prentice-Hall, Englewood Cliffs, NJ.
- GRIFFITHS, T.Vernon. (1941). An Experiment in School Music Making. New Zealand Council for Educational Research, Wellington.
- GRINDER, Robert E. (1978). Adolescence. 2nd ed. J. Wiley and Sons, New York.
- GUILFORD, J.P. and FRUCHTER, B. (1973). Fundamental Statistics in Psychology and Education. 5th ed. McGraw-Hill, New York.
- GUTTMAN, Louis. (1944). "A Basis for Scaling Qualitative Data". American Sociological Review. Vol. 9, 1944, 139-150.
- HAACK, Paul. (1990). "Beyond Objectivity: The Feeling factor in Listening". Music Educators Journal. Vol.77, No. 4, December 1990, pp.28-32.
- HEVNER, Kate. (1934). "Appreciation of Music and Tests for the Appreciation of Music". Studies in Appreciation of Art. Vol.4, No. 6 of University of Oregon Publications. Eugene: University of Oregon.
- HEVNER, Kate. (1931). "A Study of Tests for Appreciation of Music". Journal of Applied Psychology. 1931, Vol. 15, 575-583.
- HICKOK, Dorothy, and SMITH, James A. (1974). Creative Teaching of Music in the Elementary School.
- HILL, Ralph. (c. 1945). "Musical Education in Britain" in Musical Education. Ed: H. Watkins Shaw, Hinrichsen, London.
- HILLS, P. J. (Ed) (1982). A Dictionary of Education. Routledge and Kegan Paul, London.
- HIRSCH, Paul M. (1971). "Sociological Approaches to the Pop Music Phenomenon". American Behavioral Scientist. Vol. 14, 1970-71, 371-387.
- HIVELY, Wells. (Ed). (1974). Domain Referenced Testing. Educational Technology Publications, Englewood Cliffs, NJ.
- HOFFER, Charles R. (1973). Teaching Music in the Secondary Schools. 2nd ed. Wadsworth Publishing Company, Belmont, California.

- HOPKINS, Kenneth D. and STANLEY, Julian C. (1981). Educational and Psychological Measurement and Evaluation. 6th ed. Prentice-Hall, Englewood Cliffs, NJ.
- HOLT, John. (1969). How Children Fail. Pelican Books. (Reprinted 1970-71). Copyright Pitman Publishing Corporation, 1964.
- ILLICH, Ivan. (1971). Deschooling Society. Calder and Boyors, London.
- ISAAC, S. and MICHAEL, W.B. (1982). Handbook in Research and Evaluation. EDITS Publishers, San Diego.
- JAHODA, M. and WARREN, N. (1965). "The Myths of Youth". Sociology of Education. Vol. 38, 1965, 138-149.
- JANSEN, G.W. (1966). The History of School Music in New Zealand. Unpublished MA Thesis, Victoria University of Wellington.
- JANSEN, Guy. (Ed). (1969). Looking at Music Education: A Symposium by Members of the Music Education Major Course, Secondary Division. Christchurch Teachers' College.
- JENKINS, Iredell. (1971). "Aesthetic Education and Moral Refinement" in Aesthetics and Problems of Education. Ed: Ralph A. Smith, University of Illinois Press, Urbana.
- KAPLAN, M. (1966). Foundations and Frontiers in Music Education. Holt, Rinehart and Winston, New York.
- KLEINIG, John. (1982). Philosophical Issues in Education. Croom Helm, London.
- KRATHWOHL, D.R., BLOOM, B.S. and MASIA, B.B. (1964). Taxonomy of Educational Objectives, The Classification of Educational Goals. Handbook II: Affective Domain. David McKay Company, New York.
- KYME, George H. (1956). "Are Musical Tastes Indicative of Musical Capacity?". Journal of Research in Music Education. Vol. 4, Spr. 1956, No. 1, 44-51.
- LASKER, Henry. (1971). Teaching Creative Music in Today's Secondary School. Allyn and Bacon, Boston.
- LEACH, Chris. (1979). Introduction to Statistics: A Nonparametric Approach for the Social Sciences. Wiley & Sons, Chichester.
- LEEDER, J.A., and HAYNIE, W.S. (1965). Music Education in the High School. Prentice-Hall, Englewood Cliffs, NJ.
- LEHMAN, Paul R. (1968). Tests and Measurements in Music. Prentice-Hall, Englewood Cliffs, NJ.
- LEHMAN, Paul R. (1969). "A Selected Bibliography of Works on Music Testing". Journal of Research in Music Education. Vol. 17, 1969, 427-442.
- LEMON, N. (1973). Attitudes and their Measurement. Batsford, London.
- LEONHARD, Charles, and COLWELL, Richard J. (1976). "Research in Music Education". Bulletin of the Council for Research in Music Education. No. 49, Winter 1976, 1-30.

- LEONHARD, Charles, and HOUSE, Robert J. (1959) (1972). Foundations and Principles of Music Education. McGraw-Hill, New York.
- LETTS, Richard. (1973). "Musical Preferences and Aesthetic Discrimination". International Society for Music Education Yearbook 1. 1973, 122-130.
- LEWY, Arie. (1968). "The Empirical Validity of Major Properties of a Taxonomy of Affective Educational Objectives". The Journal of Experimental Education. Vol. 36, No. 3, Spring 1968, 70-77.
- LEWY, Arie. (1971). "Affective Outcomes of Musical Education". Journal of Research in Music Education. Vol. 19, 1971, 361-365.
- LEWY, Arie. (1977). Planning the School Curriculum. Fundamentals of Educational Planning No. 23. UNESCO, Paris.
- LIFTON, Walter M. (1961). "The Development of a Music Reaction Test to Measure Affective and Aesthetic Sensitivity". Journal of Research in Music Education. Vol. 9, Fall 1961, No. 2, 157-165.
- LIKERT, R., ROSLOW, S., and MURPHY, G. (1934). "A Simple and Reliable Method of Scoring the Thurstone Attitude Scales". Journal of Social Psychology. Vol. 5, 1934, 228-238.
- LINDEMAN, Richard H., and MERENDA, Peter F. (1979). Educational Measurement. 2nd ed. Scott, Foresman and Co., Illinois.
- LONG, L.R. (1948). "A New Type of Music Interest Scale". Etude. 1948, Vol. 66, 18 and 52-53.
- LONG, Newell H. (1972). "Music Discrimination Tests - Their Construction, Assumptions and Uses". The Australian Journal of Music Education. 11: 21-25.
- LUNDIN, Robert W. (1967). An Objective Psychology of Music. 2nd ed. The Ronald Press Company, New York.
- MCDONALD, Dorothy T. and SIMONS, Gene M. (1989). Musical Growth and Development: Birth Through Six. Schirmer Books, New York.
- MACINTOSH, H.G. and MORRISON, R.B. (1969). Objective Testing. Unibooks, University of London Press.
- MADSEN, C.K. and MADSEN, C.H. Jr (1978). Experimental Research in Music. Contemporary Publishing Company.
- MADSEN, C.K., GREER, R.G., and MADSEN, C.H. (Eds). (1975). Research in Music Behavior - Modifying Music Behavior in the Classroom. Teachers College Press, New York.
- MANNING, R.J. (c. 1945). "Musical Appreciation" in Musical Education. Ed: H. Watkins Shaw. Hinrichson, London.
- MARK, Michael L. (1978). Contemporary Music Education. Schirmer Books, New York.
- MARK, Michael L. (1982). Source Readings in Music Education History. Schirmer Books, New York.

- MARK, Michael L. (1982). "The Evolution of Music Education Philosophy from Utilitarian to Aesthetic". Journal of Research in Music Education. Vol. 30, Spr. 1982, 15-20.
- MASON, H.G.R. (1980). "Some Broad Tendencies" in Schools in New Zealand Society - A Book of Readings. Eds: Graham H. Robinson and Brian T. O'Rourke. Longman Paul, Auckland.
- MATTHEWS, John C. (1985). Examinations, A Commentary. George Allen and Unwin, London.
- MEEKE, Eunice Boardman, and RINEHART, Carroll. (Eds). (1975). Individualized Instruction in Music. Music Educators National Conference. Reston, Virginia.
- MEEKER, David L. (1971). "Measuring Attitude and Value Changes in Selected Humanities and Human Relations Programs". Journal of Research in Music Education. Vol. 19, 1971, 467-473.
- METCALFE, Marion. (1987). "Towards the Condition of Music" in Living Powers: The Arts in Education. Ed: Peter Abbs. The Falmer Press, London.
- METFESSEL, Newton S., MICHAEL, William B., and KIRSSNER, Donald A. (1969). "Instrumentation of Bloom's and Krathwohl's Taxonomies for the Writing of Educational Objectives". Psychology in the Schools. 1969, 6(3), 227-331.
- MILLER, Delbert C. (1970). Handbook of Research Design and Social Measurement. 2nd ed. David McKay Co.
- MOORE, J.W. (1976). "Effect of Differential Teaching Techniques on Achievement, Attitude, and Teaching Skills. Journal of Research in Music Education. 24 (3), pp.129-141.
- MORGAN, R.V.D. and MORGAN, H.N. (1960). Music Education in Action. Enlarged Edition. Neil A. Kjos Music Co., Park Ridge, Illinois.
- MUELLER, Kate Hevner. (1956). "Studies in Music Appreciation". Journal of Research in Music Education. Vol. 4, Spr. 1956, No. 1, 3-25.
- MURSELL, James L. (1937). The Psychology of Music. W.W. Norton and Company, New York.
- MURSELL, James. (1956). Music Education: Principles and Programs. Silver Burdett Co., New York.
- MUSIC EDUCATORS NATIONAL CONFERENCE NATIONAL COMMISSION ON INSTRUCTION. (1974). The School Music Program: Description and Standards. MENC.
- MUSIC EDUCATORS NATIONAL CONFERENCE. (1981). Documentary Report of the Ann Arbor Symposium - Applications of Psychology to the Teaching and Learning of Music. MENC.
- MUSIC EDUCATORS NATIONAL CONFERENCE. (1971). Toward An Aesthetic Education. MENC.

- NELSON, David J. (1983). "The Cognitive-Affective Dualism of Music Learning". Psychology of Music. Vol 11, No. 2, 67-72.
- NEW ZEALAND DEPARTMENT OF EDUCATION. (1959). The Post Primary School Curriculum. (The Thomas Report). Wellington.
- NEW ZEALAND DEPARTMENT OF EDUCATION. (1971). The Secondary School Curriculum: Some Issues and Prospects. Wellington.
- NEW ZEALAND DEPARTMENT OF EDUCATION. (1976). Towards Partnership. Report of the Committee on Secondary Education. Wellington.
- NEW ZEALAND DEPARTMENT OF EDUCATION. (1987). Draft Syllabus for Music Education - Early Childhood to Form Seven. Wellington.
- NEW ZEALAND DEPARTMENT OF EDUCATION. (1989). Syllabus for Schools: Music Education - Early Childhood to Form Seven. Wellington.
- NEW ZEALAND POST PRIMARY TEACHERS ASSOCIATION, (1969). Education in Change. Report of the Curriculum Review Group. Longman Paul, Auckland.
- NOBLE, Robert F. (1982). "A Multivariate Analysis of Factors in the Backgrounds of Wyoming Adults Related to their Attitudinal Levels Concerning Music" in A Cross-Section of Research in Music Education. Ed: Stephen H. Barnes, 228-247. Also published in the Bulletin of the Council for Research in Music Education. No. 48, Fall, 1976, 21-35.
- NOLIN, Wallace H. (1973). "Attitudinal Growth patterns Towards Elementary School Music Experiences". Journal of Research in Music Education. Vol. 21, 1973, 123-134.
- NORTHCOTE, Sydney. (c. 1945). "Music and the Adolescent" in Musical Education. Ed: H. Watkins Shaw, Hinrichsen, London.
- NSSE. (1958). Basic Concepts in Music Education. The Fifty-Seventh Yearbook of the National Society for the Study of Education. Illinois.
- OSBORN, Wendell L. (1966). A Study to Explore New Methods of Identifying and Measuring Musical Talent. Austin: University of Texas.
- PAYNTER, John. (1972). Hear and Now - An Introduction to Modern Music in Schools. Universal Editions, London.
- PAYNTER, John. (1982). Music in the Secondary School Curriculum. Cambridge University Press for the Schools Council.
- PHELPS, Roger P. (1986). A Guide to Research in Music Education. 3rd ed. The Scarecrow Press, Metuchen.
- PLUMMERIDGE, Charles. (1981). "Music in a Liberal Education" in Issues in Music Education. Bedford Way Papers 3, University of London Institute of Education.
- POGNOWSKI, Lenore M. (1985). "Attitude Assessment of Upper Elementary Students in a Process-Oriented Music Curriculum". Journal of Research in Music Education. 1985, Vol. 33, No. 4, 247-257.

- PRICE, Harry E. and SWANSON, Pamela. (1990). "Changes in Musical Attitudes Opinions, and Knowledge of Musical Appreciation Students". Journal of Research in Music Education. Vol. 38, No. 1, Spring 1990, pp.39-48.
- PRICE, Harry E. (1986). "Proposed Glossary for Use in Affective Response Literature in Music". Journal of Research in Music Education. Fall 1986, 151-159.
- PURCELL, John Leslie. (1968). Developing an Attitude Scale Using the Taxonomy of Educational Objectives: Affective Domain. Unpublished Ed.D. Thesis, Washington State University.
- RADOCY, Rudolf E. and BOYLE, J. David. (1979). Psychological Foundations of Musical Behavior. Charles Thomas Publisher, USA.
- RADOCY, Rudolf E. (1986). "On Quantifying the Uncountable in Musical Behavior". Bulletin of the Council for Research in Music Education. No. 88, Summer 1986, 22-31.
- REGELSKI, Thomas A. (1975). Principles and Problems of Music Education. Prentice-Hall, Englewood Cliffs, NJ.
- REGELSKI, Thomas A. (1981). Teaching General Music: Action Learning for Middle and Secondary Schools. Schirmer Books.
- REIMER, Bennett. (1965). "Effects of Music Education: Implications from a Review of Research". Journal of Research in Music Education. Vol. 13, 1965, 147-158.
- REIMER, Bennett. (1966). "Curriculum Reform and the Junior High General Music Class". Music Educators Journal. Vol. 53/2, Oct 1966.
- REIMER, Bennett. (1970). A Philosophy of Music Education. Prentice-Hall, Englewood Cliffs, NJ.
- REIMER, Bennett. (1971). "Building Curricula for Music Education" in Aesthetics and Problems of Education. Ed: Ralph A. Smith. University of Illinois Press, Urbana.
- REIMER, Bennett. (1986). "Towards the Transformation of Music Education". New Directions In Music Education: Yearbook 1985. New Zealand Society for Music Education (Inc). 17-31.
- RENWICK, W.L. (1986). "Opening Address". New Directions in Music Education: Yearbook 1985. New Zealand Society for Music Education (Inc). 8-15.
- REPORT OF THE COMMISSION ON EDUCATION IN NEW ZEALAND. (1962). (Currie Commission). Wellington.
- REPORT OF THE COMMITTEE TO STUDY THE NEEDS OF MUSIC TEACHING IN NEW ZEALAND. (1980). (The Ritchie Report). To the Queen Elizabeth II Arts Council. June, 1980.
- RINGNESS, Thomas, A. (1975). The Affective Domain in Education. Little, Brown & Co.
- ROBINSON, Graham H. and O'ROURKE, Brian T. (Eds). (1980). Schools in New Zealand Society - A Book of Readings. Longman Paul, Auckland.
- ROWNTREE, Derek. (1987). Assessing Students: How Shall We Know Them?. Harper and Row, London.

- ROWNTREE, Derek. (1981). A Dictionary of Education. Barnes & Noble Books, Ottawa, NJ.
- SCHOEN, Max. (1925). "Tests of Musical Feeling and Musical Understanding". Journal of Comparative Psychology. V (1925), 31-52.
- SCHWADRON, A.A. (1965). "On Relativism and Music Education". Journal of Research in Music Education. Vol. 13, 1965, No. 3, 131-135.
- SCHWADRON, A.A. (1967). Aesthetics: Dimensions for Music Education. Music Educators National Conference, Washington DC.
- SCHWADRON, A.A. (1982). "Philosophy in Music Education: Pure or Applied Research?" in A Cross-Section of Research in Music Education. Ed: Stephen H. Barnes. 275-284.
- SEASHORE, Robert H. and HEVNER, Kate. (1933). "A Time-Saving Device for the Construction of Attitude Scales". Journal of Social Psychology. Vol. 4, 1933, 366-372.
- SELL, David, F. (1986). "Five Englishmen - The British Influence on Music Education in New Zealand Between the Wars". The Canterbury Series: Studies in Music Education. No. 1, Ed: David Sell. University of Canterbury. 3-23.
- SHALLCRASS, Jack. (1980). "Schooling and Society" in Schools in New Zealand Society - A Book of Readings. Eds: Graham H. Robinson and Brian T. O'Rourke. Longman Paul, Auckland.
- SHAW, Carl N. and TOMCALA, Maryjane. (1976). "A Music Attitude Scale for Use with Upper Elementary School Children". Journal of Research in Music Education. Vol. 24, 1976, 73-80.
- SHAW, H. Watkins. (Ed). (c. 1945). Musical Education. Hinrichsen, London.
- SHAW, M.E. and WRIGHT, J.M. (1967). Scales for the Measurement of Attitudes. McGraw-Hill, New York.
- SHORT, Nicki A. (1986). "The Adolescent Sub-Cult - An Issue For Music Education". The Canterbury Series: Studies in Music Education. No. 1, Ed: David Sell. University of Canterbury. 25-33.
- SHORT, Nicki A. (1986). An Approach to General Music Education in New Zealand Secondary Schools. Unpublished Honours dissertation. University of Canterbury.
- SHUTER, Rosamund. (1968). The Psychology of Musical Ability. Methuen and Co., London.
- SIDNELL, Robert. (1973). Building Instructional Programs in Music Education. Prentice-Hall, Englewood Cliffs, NJ.
- SKINNER, B.F. (1974). About Behaviorism. Alfred A. Knopf, New York.
- SMALL, Christopher. (1977). Music-Society-Education. John Calder, London.
- SMITH, B.P.F. (1976). "Aesthetic Education" in Directions in New Zealand Secondary Education. Eds: J.A. Codd and G.L. Hermansson, Auckland.

- SMITH, R.A. and SMITH, C.M. (1971). "Justifying Aesthetic Education" in Aesthetics and Problems of Education. Ed: R.A. Smith, University of Illinois Press, Urbana.
- SMITH, Ralph A. (Ed). (1971). Aesthetics and Problems of Education. University of Illinois Press, Urbana.
- SNYDER, Bensen R. (1973). The Hidden Curriculum. MIT Press, Cambridge, Mass. Copyright 1970 by Bensen R. Snyder.
- SOMMER, Robert and SOMMER, Barbara B. (Eds.). (1986). A Practical Guide to Behavioral Research: Tools and Techniques. 2nd ed. Oxford University Press, New York.
- SONENFIELD, Irwin. (1973). "The Mystical Rite of Youth Culture: Search and Celebration in Popular Music". Music Educators Journal. 59, 6, Feb 1973, 26-31.
- SPENCE, Janet T., COTTON, John W., UNDERWOOD, Benton J., and DUNCAN, Carl P. (1983). Elementary Statistics. 4th ed. Prentice-Hall, Englewood Cliffs, New Jersey.
- SPRADLIN, J.E. (1976). "Music For Behavior Analysis". Contemporary Psychology. 21 (8), 594-595.
- STOUFFER, S.A., GUTTMAN, L., SUCHMAN, E.A., LAZARSELD, P.F., STAR, S.A., and CLAUSEN, J.A. (1950). Measurement and Prediction, Vol. IV. Princeton University Press, NJ.
- STRINGHAM, Edwin J. (1926). "Jazz - An Educational Problem". Musical Quarterly. Vol. 12, 1926, 190-195.
- SUR, W.R. and SCHULLER, C.F. (1966). Music Education for Teenagers. 2nd ed. Harper and Row, New York.
- SWANWICK, Keith. (1979). A Basis for Music Education. NFER Publishing Company, Berks.
- SWANWICK, Keith. (1988). Music, Mind and Education. Routledge, London.
- SWANWICK, Keith and TAYLOR, Dorothy. (1982). Discovering Music: Developing the Music Curriculum in Secondary Schools. Batsford, London.
- TAIT, M.J. and HAACK, Paul. (1984). Principles and Processes of Music Education. New Perspectives. Teachers College Press, New York.
- TAIT, M.J. (1970). Music Education in New Zealand. Music Teachers Registration Board, [Hamilton].
- TAYLER, E. Douglas. (1927). A Scheme of School Music Related to Human Life. Wellington.
- TAYLOR, Carwin, H. (1963). "Techniques for the Evaluation of Musical Status". Journal of Research in Music Education. Vol. 11, 1963, 55-62.
- TAYLOR, Michael. (1988). "Music Profiles - A Pilot Scheme". British Journal of Music Education. Vol. 3, No. 1, 1988, 19-33.

- THORN, James. (1952). Peter Fraser, New Zealand's Wartime Prime Minister. Odhams Press, London.
- THORNDIKE, Edward L. and HAGEN, Elizabeth P. (1977). Measurement and Evaluation in Psychology and Education. 4th ed. John Wiley and Sons.
- THORNDIKE, Edward L. BREGMAN, E.D., COBB, M.V. AND WOODYARD, E. (1927). The Measurement of Intelligence. Columbia University, New York.
- THURSTONE, L.L. (1959). The Measurement of Values. University of Chicago Press.
- TUNKS, Thomas, W. (1973). "An Application of Guttman Facet Theory To Attitude Scale Construction in Music". Bulletin of the Council for Research in Music Education. No. 33, Summer 1973, 47-53.
- TUNKS, Thomas, W. (1987). "Evaluation in Music Education: The Value of Measurement/The Measurement of Value". Bulletin of the Council for Research in Music Education. No. 90, Winter 1987, 53-59.
- TYLER, R.W. (1973). "Assessing Educational Achievement in the Affective Domain". NCME Measurement in Education. Vol. 4. No. 3, Spring 1973.
- UDINSKY, B.F., OSTERLIND, S.J. and LYNCH, S.W. (1981). Evaluation Resource Handbook: Gathering, Analyzing, Reporting Data. Edits Publishers, San Diego, California.
- VULLIAMY, Graham, and LEE, Ed. (1980). Pop Music in School. New Edition. Cambridge University Press.
- WALDEN MILLS, J. (1962). "Music in New Zealand Schools". Education. Vol. 11, 1962, 7-11.
- WAPNICK, Joel. (1976). "A Review of Research on Attitude and Preference". Bulletin of the Council for Research in Music Education. Vol. 48, Fall 1976, 1-20.
- WHEELER, D.K. (1971). "Aesthetic Education and Curriculum" in Aesthetics and Problems of Education. Ed: Ralph A. Smith. University of Illinois Press, Urbana.
- WHEELWRIGHT, Phillip Ellis. (1951). Aristotle. Odyssey Press, New York.
- WHYBREW, William, E. (1962). Measurement and Evaluation in Music. WMC Brown, Dubuque, Iowa.
- WHYBREW, William, E. (1973). "Research in Evaluation in Music Education". Bulletin of the Council for Research in Music Education. No. 35, Winter 1973, 9-17.
- WILLIAMS, Robert, O. (1972). "Effects of Musical Aptitude, Instruction and Social Status On Attitudes Toward Music". Journal of Research in Music Education. Vol. 20, Fall 1972, 362-369.
- WING, Herbert. (1948). "Tests of Musical Ability and Appreciation". The British Journal of Psychology Monography Supplements. XXVII, Cambridge University Press.
- WING, Herbert. (1960). Manual for Standardised Tests of Musical Intelligence. City of Sheffield Training College.
- WITKIN, Robert W. (1974). The Intelligence of Feeling. Heinemann, London.